The Einstein Field Equations

on semi-Riemannian manifolds, and the Schwarzschild solution

Rasmus Leijon
Abstract

Semi-Riemannian manifolds is a subject popular in physics, with applications particularly to modern gravitational theory and electrodynamics. Semi-Riemannian geometry is a branch of differential geometry, similar to Riemannian geometry. In fact, Riemannian geometry is a special case of semi-Riemannian geometry where the scalar product of nonzero vectors is only allowed to be positive.

This essay approaches the subject from a mathematical perspective, proving some of the main theorems of semi-Riemannian geometry such as the existence and uniqueness of the covariant derivative of Levi-Civita connection, and some properties of the curvature tensor.

Finally, this essay aims to deal with the physical applications of semi-Riemannian geometry. In it, two key theorems are proven - the equivalence of the Einstein field equations, the foundation of modern gravitational physics, and the Schwarzschild solution to the Einstein field equations. Examples of applications of these theorems are presented.
# Contents

1 Introduction 4

2 Introduction to Semi-Riemannian Manifolds 8  
   2.1 Preliminaries ........................................... 8  
   2.2 Geometry ............................................... 9  
   2.3 Distance ................................................ 18  
   2.4 Geodesics and parallel transport ....................... 19  
   2.5 Curvature ............................................... 22

3 Semi-Riemannian Geometry 28  
   3.1 Riemannian geometry ................................. 28  
   3.2 Lorentzian geometry .................................. 29  
   3.3 Examples ............................................... 31

4 The Einstein Field Equations 34  
   4.1 Special relativity ..................................... 34  
      4.1.1 Generalization of Newtonian concepts .......... 35  
   4.2 General relativity .................................... 36  
      4.2.1 The Einstein field equations .................. 37  
   4.3 The Schwarzschild geometry .......................... 39

5 Applications to the Schwarzschild solution 50  
   5.1 The Schwarzschild manifold .......................... 50  
   5.2 Schwarzschild geodesics .............................. 50  
   5.3 Schwarzschild black holes ............................. 52
1 Introduction

In 1905, Einstein published his *annus mirabilis* papers, among them the theory of special relativity ([1]). However, this theory had a flaw: incompatibility with gravitational fields. This led Einstein to consider the generalization of this theory. In 1908, Minkowski examined the concept of spacetime from a mathematical perspective, introducing semi-Riemannian manifolds to assist with the physical calculations ([10]). However, his original formulation of the differential geometry of these semi-Riemannian manifolds were not particularly mathematical ([10]).

However, in terms of this semi-Riemannian geometry, Einstein formulated his theory of general relativity over the next decade, culminating in the Einstein field equations (see e.g. [2] and [3]). The general theory of relativity relates the energy-momentum content of the physical universe to the curvature of the model manifold through these equations.

The first solution presented to the Einstein field equations was published by Karl Schwarzschild in 1916 ([19]), which has allowed us to make many physical predictions with increased precision. It is the unique solution for the field outside a static, spherically symmetric body (see e.g. [4]). Even more impressively, Schwarzschild was actively serving as a lieutenant in the German artillery in the first World War while solving the Einstein field equations, dying from infection only some months after publishing his paper ([6]).

Examples of the improvements to physical theory includes the perihelion advance of planets in our solar system, the deflection of light due to massive objects, the spectral shift of light due to massive objects and many more (see, e.g. [11] or [4]). From a physical perspective, this is very interesting. However, the Schwarzschild solution is not the only known solution to the Einstein field equations. Some of the famous solutions include the Kerr solution (for the spacetime surrounding a rotating mass), the Reissner-Nordström solution (for the spacetime surrounding a charged mass) and the Kerr-Newman solution (for the spacetime surrounding a charged and rotating mass) (see e.g. [5], chapter 33 in [11] and [15]).

From a mathematical perspective, the theory of general relativity is interesting chiefly due to the semi-Riemannian geometry it is formulated in. In this essay, we will pursue a deeper understanding of these semi-Riemannian manifolds and the geometries they result in. A semi-Riemannian manifold is a smooth manifold associated with a scalar product (a metric tensor) on its tangent bundle. This allows for mathematical generalizations of length, volume and curvature. For this reason, the reader should be familiar with smooth manifolds and tensor algebra, as well as real analysis (see e.g. [21], [9] and [18] for books on these subjects).

In the theory of smooth manifolds, much effort is usually put into the smooth structure of the manifold. In the theory of semi-Riemannian manifolds, focus is instead directed at the geometry of the semi-Riemannian manifold. The reader may be familiar with Riemannian geometry, a branch of differential geometry introduced by Bernhard Riemann in the middle of the 19th century ([17]). In Riemannian geometry, we study the geometry of smooth manifolds with an inner product (note the stronger requirement compared to the semi-Riemannian case).
on each tangent space of the manifold. Riemann’s formulation of Riemannian geometry is a convenient generalization of the differential geometry of surfaces in $\mathbb{R}^n$.

However, we are interested in the more general semi-Riemannian geometry, where the inner product on each tangent space of the manifold is replaced with a scalar product (no longer required to be positive definite). This may seem a trivial generalization of Riemannian geometry, but the geometries of semi-Riemannian manifolds and Riemannian manifolds are in general completely different. In semi-Riemannian geometry, we generalize many of the concepts we find useful in the differential geometry on surfaces in Euclidean space such as curvatures, straight lines and distances.

The main theorems, Theorems 4.4 (Theorem A) and 4.8 (Theorem B), of this paper are presented here, and the schematics of the proofs are given, giving an implicit insight into the content and purpose of each chapter of this essay.

**Theorem A: The Einstein Tensor**

The Einstein tensor, $G = \text{Ric} - \frac{1}{2}gS$, is a symmetric, $(0,2)$ tensor field with vanishing divergence.

To construct the proof of this theorem, we need pieces taken from all parts of the foundation of semi-Riemannian geometry.

In chapter 2, we define the metric tensor $g$ on the tangent bundle of a semi-Riemannian manifold (Definition 2.6). This tensor is a symmetric $(0,2)$ tensor field.

In section 2.2, We then prove the uniqueness and existence of the covariant derivative (Theorem 2.17) on a semi-Riemannian manifold, which is a type of connection (see, e.g. chapters 12 and 13 in [7]). In section 2.5, we define in terms of the covariant derivative the curvature tensor $\mathbf{R}$ (Definition 2.25). The (Riemann) curvature tensor $\mathbf{R}$ expresses the curvature of the semi-Riemannian manifold. From this curvature tensor, we derive the Ricci tensor (Definition 2.29), $\text{Ric}$, which is a contraction of the curvature tensor. Similarly, a contraction of the Ricci tensor gives us the scalar curvature $S$ (Definition 2.31). It is important for the proof of Theorem A that the Ricci tensor is symmetric, and we prove this later in the section (Lemma 2.30).

Finally, we use the second Bianchi identity (also derived in chapter 2) to prove that $\text{div} G = 0$. Now it is clear what we need in order to prove Theorem A.

From the Einstein tensor, we define a set of nonlinear partial differential equations solved for the geometry of the semi-Riemannian manifold $\mathfrak{M}$ called the *Einstein field equations*, given by

$$\text{Ric} - \frac{1}{2}gS = kT$$

(1.1)

where $T$ is the stress-energy tensor (Definition 4.2) which represents the energy-momentum content of the universe in the physical model. This is the main focus of this paper on semi-Riemannian manifolds and perhaps one of the main focus of the entire field of semi-Riemannian geometry.
Theorem B: The Schwarzschild Solution
Let $(\mathcal{M}, g)$ be a 4-dimensional Lorentzian manifold that is static, spherically symmetric and flat at infinity. Then in terms of the chart $(U, \phi = (ct, R, \theta, \varphi))$ the Einstein field equations (1.1) have a unique solution given by:

$$g_{\alpha\beta} = \begin{pmatrix}
(1 - \frac{r_S}{R}) & 0 & 0 & 0 \\
0 & -\left(1 - \frac{r_S}{R}\right)^{-1} & 0 & 0 \\
0 & 0 & -R^2 & 0 \\
0 & 0 & 0 & -R^2 \sin^2 \theta
\end{pmatrix}$$

where $r_S$ is a constant.

Theorem B is one solution to the Einstein field equations, defined earlier. However, we will not only need an understanding of the Einstein field equations, but also a better understanding of the requirements on the semi-Riemannian manifold. First, we shall clear up what we mean by a 4-dimensional Lorentzian manifold. It is a semi-Riemannian manifold of metric signature $(1,3)$ (Definitions 3.2, 2.7 and 2.11, respectively).

Now, we will clearly also have to consider what we mean by spherically symmetric and static. We will find a definition of spherically symmetric in more detail in e.g. Example 2.10 and section 3.3. A clearer definition will also be presented and justified in section 4.3. The term static is related to what we will call causal characterization on Lorentzian manifolds. A deeper and detailed explanation of causal characterization and Lorentzian manifolds will be given in section 3.2.

Finally, we will look at the Ricci tensor again to determine flatness at infinity, because what we mean by flat is that the Ricci tensor vanishes. Now, solving these equations will take some work, however once topological and geometrical properties of the manifold have been made rigorous, the problem is mainly calculus, and we shall solve it in great detail in section 4.3.

After these two theorems have been proven, we dedicate the final chapter, chapter 5, to geodesics (Definition 2.21) and black holes on the Schwarzschild manifold. Geodesics are, in short, the generalizations of straight lines to the curved geometry of semi-Riemannian manifolds, and as such are interesting mathematically and physically, as we shall later see. Hence, chapter 5 gives two examples; geodesics of particles in orbit (crash- and bound orbits) around a massive object as well as mathematical considerations of the Schwarzschild black hole.

Remark. From now on we will omit the word ‘smooth’ in this paper. It is then considered implied that all functions, manifolds, structures, curves, and so on, are smooth, unless otherwise noted or inferred from context.
2 Introduction to Semi-Riemannian Manifolds

In this chapter, we consider generalizations of geometrical concepts that we are at least intuitively familiar with in Euclidean space. We will talk about geometry and metrics. In particular, the question we will ask ourselves is: how do we measure lengths of curves and vectors?

This concept of metrics in geometry is not the same as the metric one would consider in topology. In fact, most of the functions we will talk about in this chapter are not functions on a manifold itself, but on the tangent bundle or sections of the tangent bundle associated with a given manifold. We will distinguish between the two when we feel it is important, but most of the time it is clear from context what we mean.

Furthermore, we will discuss curvature and parallel transport. Two moderately extensive examples are given in section 4.3, and the curious reader may skip ahead.

It may be necessary to explain why we are starting with semi-Riemannian manifolds instead of Riemannian manifolds, since semi-Riemannian manifolds are in one sense an extension of Riemannian manifolds. From our perspective, however, Riemannian manifolds are simply a special case of semi-Riemannian manifolds, and will be regarded as such throughout this paper, and Riemannian manifolds will be explored in the section 4.1. Similarly, Lorentzian manifolds are simply a special case of semi-Riemannian manifolds from our perspective, albeit the most interesting special case due to the practical applications of these manifolds. Therefore we will deal with them in the following chapters.

2.1 Preliminaries

Before we start discussing geometry, we have a few preliminary definitions to take care of. The reader should already be familiar with the topology of manifolds, however we shall define some of the central mathematical objects that we will be using. This is for added clarity, since the reader may be used to some other definition.

Definition 2.1 (Chart). An $n$-dimensional chart on a space $\mathcal{M}$ is a pair $(\mathcal{U}, \phi)$ with $\mathcal{U} \subseteq \mathcal{M}$, and with a coordinate map $\phi: \mathcal{U} \to \phi(\mathcal{U}) \subseteq \mathbb{R}^n$ that is a bijection from $\mathcal{U}$ onto some part of $\mathbb{R}^n$.

We will talk about charts as coordinate systems on $\mathcal{U}$, with coordinates $\phi = (u^1, u^2, ..., u^n)$. Similarly, we will think of atlases as a sort of collection of mutually overlapping charts on a manifold that allows us to construct a coordinate system for the manifold.

Definition 2.2 (Atlas). An $n$-dimensional atlas on a space $\mathcal{M}$ is a collection of charts $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$, that satisfies $\mathcal{M} = \bigcup \mathcal{U}_\alpha$. For every pair of charts, $(\mathcal{U}_\alpha, \phi_\alpha)$, $(\mathcal{U}_\beta, \phi_\beta)$, it is required that $\phi_\alpha \circ \phi_\beta^{-1}$ and $\phi_\beta \circ \phi_\alpha^{-1}$ are smooth functions.

We will at several points in this paper return to the topology of manifolds. Therefore we need to define the structure on a manifold.
Definition 2.3 (Structure). A structure on a set $\mathcal{M}$ is an equivalence class of atlases: $\{A_\beta : A_\beta \cup A_\lambda \text{ is an atlas for all } \beta, \lambda\}$.

Definition 2.4 (Maximal atlas). A maximal atlas $\mathcal{A}$ is the union of all elements in the structure, i.e. $\mathcal{A} = \bigcup_{\alpha} A_\alpha$.

Now we are ready to introduce the manifold.

Definition 2.5 (Manifold). An $n$-dimensional manifold is a set with a maximal atlas such that the topology induced by the maximal atlas is second countable and Hausdorff.

The maximal atlas through its associated charts provides a basis for a topology on the manifold $\mathcal{M}$, and it is this we mean by the topology being induced by the maximal atlas. Note also that the topology induced by the atlas makes $\mathcal{M}$ a topological manifold. It is clear that this manifold is locally homeomorphic to $\mathbb{R}^n$. Finally, we are ready to focus on the geometry of manifolds.

2.2 Geometry

The metric tensor is what turns a manifold into a semi-Riemannian manifold, and allows for all the interesting and sometimes counter-intuitive properties and applications of different semi-Riemannian manifolds that is the focus of efforts with these spaces. Again, note that this is not the same thing as a topological metric (see e.g. chapters 7 and 8 in [18]).

Definition 2.6 (Metric tensor). Let $\mathcal{M}$ be a manifold and $T\mathcal{M}$ the tangent bundle of the manifold. The metric tensor, $g$, on $\mathcal{M}$ is a $(0,2)$ tensor field on $T\mathcal{M}$, $g: T\mathcal{M} \times T\mathcal{M} \to \mathbb{R}$, such that it assigns to each tangent space a scalar product, $g_p = \langle X_p, Y_p \rangle$.

This $g_p$ is required to vary smoothly over all points $p \in \mathcal{M}$. The metric tensor is required to have the following properties for any two vector fields $X$, $Y$ (or indeed vectors):

i. Symmetric: $g(X,Y) = g(Y,X)$,

ii. Nondegenerate: The set of functions $\{f_c : Y \mapsto g(X_c, Y)\}$, for $X_c$ a fixed vector field, is for all $c$ not identically zero, provided $X_c \neq 0$.

It should be noted that this is not the definition of an inner product space. An inner product $g_p$ is required to be positive definite, i.e. $g_p(X, Y) > 0$, $X, Y \neq 0$. However, this is not the case for this metric tensor, which is only required to be nondegenerate (see e.g. chapter 3 in [16]).

Had we defined this metric tensor as an inner product space on each tangent space, we would have defined the metric tensor of a Riemannian manifold. Instead we keep this weaker definition and delve into the differences between Riemannian and semi-Riemannian further in chapter 4.

However, this is not the most general definition possible, and in fact there are applications that make use of a nonsymmetric metric tensor - which of course is not really a metric tensor of a semi-Riemannian manifold! This definition does not require $g(X,Y) = g(Y,X)$, which may seem very odd, but this definition...
has its uses, for example in the nonsymmetric gravitational theory (see e.g. [12]).

Now we define our notation for the metric tensor. Keep in mind that it is a type of generalization of the inner product in Euclidean space, and is therefore used to measure, for example, the length of vectors. However, note that it is not required to be positive-definite, and therefore allows for semi-Riemannian manifolds to be a type of manifold with very peculiar properties.

Let us introduce the coordinate basis of some coordinate system on \( \mathcal{M} \). Then in terms of the chart \((\mathcal{U}, \phi = (u^1, u^2, ..., u^n))\), we denote the coordinate basis by the set \( \partial_1, \partial_2, ..., \partial_n \), where of course we use the notation \( \partial_\alpha = \partial / \partial u^\alpha \).

We denote the components of \( g \) relative to some coordinate basis by \( g_{\alpha\beta} = \langle \partial_\alpha, \partial_\beta \rangle \),

and the \( n \times n \) matrix of the components of \( g \) relative to an \( n \)-dimensional chart on its semi-Riemannian manifold by \( [g_{\alpha\beta}] \). Denote the inverse of this matrix by \( g^{-1} \), with components \( g^{\alpha\beta} \). This inverse must exist due to the nondegenerate nature of the metric tensor.

Let \( X, Y \) be two vector fields \( X = \sum_\alpha X^\alpha \partial_\alpha \) and \( Y = \sum_\alpha Y^\alpha \partial_\alpha \). Here \( X^\alpha \) and \( Y^\alpha \) are the real-valued functions called the components of the vector fields \( X \) and \( Y \). Since \( g \) is bilinear, \( g(X, Y) \) can be written in terms of the product of its components:

\[
g(X, Y) = \langle X, Y \rangle = \sum_\alpha \sum_\beta X^\alpha Y^\beta \langle \partial_\alpha, \partial_\beta \rangle = \sum_\alpha \sum_\beta g_{\alpha\beta} X^\alpha Y^\beta.
\]

**Definition 2.7** (Semi-Riemannian manifold). A manifold \( \mathcal{M} \) with an associated metric tensor \( g \) forms a pair \((\mathcal{M}, g)\). This pair is called a semi-Riemannian manifold.

This has defined the most general space we will be looking at in this paper, namely the semi-Riemannian manifold. Usually we will denote a semi-Riemannian manifold with metric tensor \( g \) by \( \mathcal{M} \).

We are now interested in looking at some examples of semi-Riemannian manifolds. We will look at simple but useful examples. First, we study Euclidean space, which the reader is probably intuitively familiar with (see e.g. chapter 1 in [21] for a more thorough understanding).

**Example 2.8** (Euclidean space). Let us denote the \( n \)-dimensional manifold of Euclidean space by \( \mathbb{R}^n \), and its metric tensor by \( g \). The cartesian coordinates \((\mathcal{U}, \phi = (x^1, x^2, ..., x^n))\) forms a structure for \( \mathbb{R}^n \) by \( \mathcal{U} = \mathbb{R}^n \), which means the coordinates \((\mathcal{U}, \phi)\) are global, which is rare in general. For this semi-Riemannian manifold, the metric tensor is given by:

\[
\langle \partial_\alpha, \partial_\beta \rangle = g_{\alpha\beta} = \begin{cases} 
1 & \text{for all } \alpha = \beta, \\
0 & \text{otherwise}.
\end{cases}
\]

Euclidean space is perhaps the most common space used, and therefore it is useful to think of it as a special case of semi-Riemannian manifold. Since the metric tensor is positive definite, this is in fact also a Riemannian manifold. □
Now we will look at a generalization of Euclidean space, which is called semi-Euclidean space (see e.g. chapter 3 in [16] for more detail). It is useful to understand the difference between Riemannian manifolds and semi-Riemannian manifolds as somewhat analogous to the difference between Euclidean space and semi-Euclidean space.

**Example 2.9** (Semi-Euclidean space). Let us denote the \((n + \nu)\)-dimensional manifold of semi-Euclidean space by \(\mathbb{R}^n_\nu\), and its metric tensor by \(g\). In terms of the cartesian coordinates \((U, \phi = (x^1, x^2, \ldots, x^n))\), the structure of \(\mathbb{R}^n_\nu\) is the same as for the Euclidean space \(\mathbb{R}^{(n+\nu)}\), which means that \((U, \phi = (x^1, x^2, \ldots, x^n))\) is a global coordinate system. For \(\mathbb{R}^n_\nu\), the metric tensor is given by:

\[
\langle \partial_\alpha, \partial_\beta \rangle = g_{\alpha\beta} = \begin{cases} 
1 & \text{for all } \alpha = \beta \leq \nu, \\
-1 & \text{for all } \alpha = \beta > \nu, \\
0 & \text{otherwise}.
\end{cases}
\]

We shall think of semi-Euclidean space as a model space for a semi-Riemannian manifold. Notice that clearly in this metric, the scalar product of two vectors can be negative.

As a last example of a semi-Riemannian manifold, we take the 2-sphere \(S^2\) (for more detail, see e.g. chapter 5 in [21]).

**Example 2.10** (2-Sphere). Let us define \(S^2\) as the sphere of unit radius with associated metric tensor \(g\). Using the cartesian coordinates of Euclidean \(\mathbb{R}^3\), defined globally, we set \(S^2\) as all the points satisfying \((x^1)^2 + (x^2)^2 + (x^3)^2 = 1\).

Typically, we find an atlas on \(S^2\) by dividing the sphere into six hemispheres (six charts):

\[
\begin{align*}
\mathcal{U}_1 &= \{(x^1, x^2, x^3) : x^1 > 0\}, & \phi &= (x^2, x^3), \\
\mathcal{U}_2 &= \{(x^1, x^2, x^3) : x^1 < 0\}, & \phi &= (x^2, x^3), \\
\mathcal{U}_3 &= \{(x^1, x^2, x^3) : x^2 > 0\}, & \phi &= (x^1, x^3), \\
\mathcal{U}_4 &= \{(x^1, x^2, x^3) : x^2 < 0\}, & \phi &= (x^1, x^3), \\
\mathcal{U}_5 &= \{(x^1, x^2, x^3) : x^3 > 0\}, & \phi &= (x^1, x^2), \\
\mathcal{U}_6 &= \{(x^1, x^2, x^3) : x^3 < 0\}, & \phi &= (x^1, x^2).
\end{align*}
\]

On each one, only two coordinates are needed to describe each point. Expressed in what we will call spherical coordinates, \((\mathcal{U}, \phi = (\theta, \varphi))\). This gives us the metric tensor components

\[
\langle \partial_\alpha, \partial_\beta \rangle = g_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{bmatrix}
\]

Now we have a good idea of the structure of the 2-sphere, as well as its geometry. It is a type of Riemannian manifold as we will see in the next chapter. This 2-sphere will be examplified in section 3.3, since we will use it in the applications to gravitational theory.

We will need to define a few more properties of the metric on a semi-Riemannian manifold. Clearly, we will need a way to categorize semi-Riemannian manifolds to distinguish between manifolds with different properties. We use what is called the metric signature to do this. (see e.g. chapter 1 in [9]).
Definition 2.11 (Metric signature). The signature of a metric tensor $g$ is the number of positive ($m$) and negative ($n$) eigenvalues of the matrix $[g_{\alpha\beta}]$, usually denoted as a pair, $(m,n)$, $m,n \in \mathbb{N}$. The manifold $\mathcal{M}$ of the metric tensor $g$ is said to be of dimension $m+n$, and is said to have the metric signature $(m,n)$.

Now we introduce a brief example so that the significance of this classification is more obvious to the reader. We shall show the metric signature of the Minkowski space, which is a semi-Riemannian manifold named after the 19th century mathematician Hermann Minkowski ([10]).

Example 2.12 (Metric signature of a Minkowski space). Let $\mathcal{M}$ be a semi-Euclidean space with $n=1$ and $\nu=3$, i.e. $\mathcal{M} = \mathbb{R}^4_1$. Then we call $\mathcal{M}$ the Minkowski space. Then it has a metric tensor that is usually called $\eta$. For cartesian coordinates, the chart $(\mathcal{U}, \phi = (x^1, x^2, x^3, x^4))$, where once again $\mathcal{U} = \mathbb{R}^{(n+\nu)}$ is global. Then the components of $\eta$ are given by

$$\langle \partial_\alpha, \partial_\beta \rangle = \eta_{\alpha\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

The metric tensor clearly has one positive eigenvalue and three negative eigenvalues and therefore a metric signature of $(1,3)$.

The Minkowski space is a semi-Riemannian manifold that is used particularly in the special theory of relativity. We will talk about it in a little more detail in section 4.1.

Furthermore, we will need to introduce a useful tool when dealing with distances on a manifold. This is called the line element, and it can be thought of as describing an infinitesimal distance on a manifold.

Definition 2.13 (Line element on a manifold). Let $\mathcal{M}$ be a semi-Riemannian manifold with associated metric tensor $g$. Then, in terms of a coordinate chart $(\mathcal{U}, \phi = (u^1, u^2, ..., u^n))$, the line element $q$ is defined as a quadratic form of $g_{\alpha\beta}$, such that $q(X) = \langle X, X \rangle$. Since

$$q(X) = \sum_{\alpha} \sum_{\beta} g_{\alpha\beta} X^\alpha X^\beta = \sum_{\alpha} \sum_{\beta} g_{\alpha\beta} du^\alpha(X) du^\beta(X),$$

we write $q$ as the product

$$q = \sum_{\alpha} \sum_{\beta} g_{\alpha\beta} du^\alpha du^\beta,$$

This definition is used for example in chapter 3 in [16].

From these definitions, particularly the definition of a metric tensor, we can find new properties of the tangent bundles of our manifolds. Remembering that $[X,Y]$ is the Lie bracket of vector fields, we can define the covariant derivative, which is a connection on the tangent bundle (see, e.g. chapter 5 in [7] for the definition of Lie bracket).

12
Definition 2.14 (Covariant derivative). Let \( \mathcal{M} \) be a semi-Riemannian manifold with tangent bundle \( T\mathcal{M} \). Let \( X \) and \( Y \) be vector fields in the tangent bundle.

Then we define the covariant derivative relative to \( X \), \( \nabla_X \), as a function \( \nabla_X Y : \mathfrak{X}(U) \times \mathfrak{X}(U) \to \mathfrak{X}(U) \) with the following properties:

i. The covariant derivative of a real-valued function coincides with the directional derivative of the function: \( \nabla_X f = Xf \),

ii. It is linear in the first variable and additive in the second variable:

\[
\nabla_{fX + gY}(Z + W) = f\nabla_X Z + g\nabla_Y Z + f\nabla_X W + g\nabla_Y W,
\]

iii. It is Leibnizian in the second variable:

\[
\nabla_X fY = f\nabla_X Y + Y\nabla_X f = f\nabla_X Y + YXf,
\]

iv. It is torsion-free: \( \nabla_X Y - \nabla_Y X = [X,Y] \),

v. It is compatible with the metric: \( \langle X, \nabla_Y Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \),

for \( X,Y,Z,W \in \mathfrak{X}(U) \) and \( f,g \in C^\infty(U) \).

This definition of the covariant derivative is often called the Levi-Civita connection, and is a special type of connection. Many other authors make use of the more general object of connections, usually defined as a function with the properties i) through iii) (see, e.g. chapter 3 in [16] or chapter 12 in [7]).

Since we do not take interest in any connection other than the covariant derivative, this definition is sufficient.

This definition of the covariant derivative will lead to a special type of function, namely the covariant derivative of a coordinate vector relative to another coordinate vector. Even though this may seem like an innocuous enough function, we will soon see that it is useful in describing curves on a manifold.

Since we can write a vector field \( X \) in terms of the coordinate basis as \( X = \sum_\alpha X^\alpha \partial_\alpha \), with \( X^\alpha \in C^\infty(\mathcal{M}) \), we must surely be able to do the same thing with the covariant derivative (see e.g. chapter 4 in [7]).

Definition 2.15 (Christoffel symbols). A chart \((U, \phi)\) on a semi-Riemannian manifold \( \mathcal{M} \) has a set of Christoffel symbols, functions \( \Gamma^\lambda_{\alpha\beta} : U \to \mathbb{R} \) defined by:

\[
\nabla_{\partial_{\alpha}} \partial_{\beta} = \sum_\lambda \Gamma^\lambda_{\alpha\beta} \partial_\lambda.
\]

Lemma 2.16 (The Koszul formula). Let \( \mathcal{M} \) be a semi-Riemannian manifold with associated metric tensor \( g \). Then the covariant derivative satisfies the Koszul formula:

\[
2 \langle \nabla_X Y, Z \rangle = X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle - \langle Y, [X,Z] \rangle - \langle Z, [Y,X] \rangle - \langle X, [Z,Y] \rangle.
\] (2.1)

Proof. Inspired by the proof Theorem 11 in [16], according to property v) of the covariant derivative it is metric compatible, i.e.

\[
X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.
\]

13
Cyclically permuting this equation, we get two more equations,

\[ Y \langle Z, X \rangle = \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle, \]

\[ Z \langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle. \]

Now, from the torsion-free property of the covariant derivative, we rewrite the last term of each equation as \( \nabla_X Y = [X, Y] + \nabla_Y X. \) These three equations are then given by

\[ X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_Z X \rangle + \langle Y, [X, Z] \rangle, \quad (2.2a) \]

\[ Y \langle Z, X \rangle = \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_X Y \rangle + \langle Z, [Y, X] \rangle, \quad (2.2b) \]

\[ Z \langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Y Z \rangle + \langle X, [Z, Y] \rangle. \quad (2.2c) \]

Remembering that the metric tensor is symmetric, we can simplify the right-hand side of (2.2a) + (2.2b) - (2.2c), giving us:

\[ 2 \langle \nabla_X Y, Z \rangle - 2 \langle \nabla_Y X, Z \rangle = 0, \]

which becomes:

\[ 2 \langle \nabla_X Y, Z \rangle = X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle \]

\[ - \langle Y, [X, Z] \rangle - \langle Z, [Y, X] \rangle - \langle X, [Z, Y] \rangle. \]

This is called the Koszul formula and it is a useful property as we shall see in the proof of Theorem 2.17.

After this technical lemma, we are ready to prove the main theorem of this section.

**Theorem 2.17** (Existence and uniqueness of the covariant derivative). Let \( M \) be a semi-Riemannian manifold with associated metric tensor \( g \). Then there exists a unique covariant derivative on \( M \).

**Proof.** Inspired by the proof of Theorem 11 in [16], we prove this theorem. First we prove the uniqueness of the covariant derivative: If the covariant derivative is not unique, then there exists another covariant derivative, say \( \nabla' \). However, note that the right-hand side of the Koszul formula (2.1) does not depend on the covariant derivative.

Therefore, we use the Koszul formula for the two operators \( \nabla \) and \( \nabla' \). This gives us

\[ 2 \langle \nabla'_X Y, Z \rangle - 2 \langle \nabla_X Y, Z \rangle = 0, \]

and

\[ \langle (\nabla'_X - \nabla_X) Y, Z \rangle = 0, \]

and obviously this proves that \( \nabla' = \nabla \).
Now we prove the existence of the covariant derivative: Since any vector can be written in terms of a coordinate basis, we may without loss of generality use the coordinate basis of some chart, \((U, \phi)\). We then know that \([\partial_\alpha, \partial_\beta] = 0\) for all \(\alpha, \beta\).

Because of this, we can rewrite the Koszul formula (2.1) in terms of the coordinate basis

\[
2\langle \nabla_{\partial_\alpha} \partial_\beta, \partial_\lambda \rangle = \partial_\alpha \langle \partial_\beta, \partial_\lambda \rangle + \partial_\beta \langle \partial_\lambda, \partial_\alpha \rangle - \partial_\lambda \langle \partial_\alpha, \partial_\beta \rangle. \tag{2.3}
\]

From the definition of the Christoffel symbols (2.15), and the definition of the metric tensor components, we can write equation (2.3) as

\[
\sum_\mu \Gamma^\mu_{\alpha\beta} g_{\mu\lambda} = \frac{1}{2} (\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\lambda\alpha} - \partial_\lambda g_{\alpha\beta}). \tag{2.4}
\]

Multiplying by the metric tensor inverse, \(g^{\lambda\nu}\) and summing over \(\lambda\) we finally arrive at,

\[
\Gamma^\nu_{\alpha\beta} = \frac{1}{2} \sum_\lambda g^{\lambda\nu} (\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\lambda\alpha} - \partial_\lambda g_{\alpha\beta}), \tag{2.5}
\]

due to the property \(\sum_\alpha g_{\alpha\beta} g^{\alpha\lambda} = \delta^\lambda_\beta\), where \(\delta^\lambda_\beta\) is the Kronecker delta function.

This clearly defines a covariant derivative on any chart \((U, \phi)\), since it fulfills all conditions for a covariant derivative.

**First we prove that property i) holds:** This property follows from the proof of property iii), which we will soon get to.

**Now we prove that property ii) holds:** We use the Koszul formula (2.1) as the basis for this proof. First we prove linearity in the first variable. As we explained previously, we can without loss of generality use \(f \partial_\alpha + h \partial_\beta\) as \(X\), \(\partial_\lambda\) as \(Y\) and \(\partial_\nu\) as \(Z\) in the Koszul formula. This gives us:

\[
2\langle \nabla_{f \partial_\alpha + h \partial_\beta} \partial_\lambda, \partial_\nu \rangle = (f \partial_\alpha + h \partial_\beta) \langle \partial_\lambda, \partial_\nu \rangle + \partial_\lambda \langle (f \partial_\alpha + h \partial_\beta), \partial_\nu \rangle
- \partial_\nu \langle (f \partial_\alpha + h \partial_\beta), \partial_\lambda \rangle - \langle \partial_\lambda, [(f \partial_\alpha + h \partial_\beta), \partial_\nu] \rangle
- \langle \partial_\nu, [\partial_\lambda, (f \partial_\alpha + h \partial_\beta)] \rangle - \langle \partial_\nu, [(f \partial_\alpha + h \partial_\beta), \partial_\lambda] \rangle. \tag{2.6}
\]

What we will prove is therefore that the right-hand side of (2.6) is equal to

\[
2f \sum_\rho \Gamma^\rho_{\alpha\lambda} g_{\rho\nu} + 2h \sum_\rho \Gamma^\rho_{\beta\lambda} g_{\rho\nu}.
\]

This is proven by rewriting the right-hand side of equation (2.6), and using the linearity property of the metric tensor, and by \([\partial_\nu, \partial_\lambda] = 0\) for all \(\nu, \lambda\), this becomes

\[
f \partial_\alpha \langle \partial_\lambda, \partial_\nu \rangle + h \partial_\beta \langle \partial_\lambda, \partial_\nu \rangle + \partial_\lambda \langle f \partial_\alpha, \partial_\nu \rangle + \partial_\lambda \langle h \partial_\beta, \partial_\nu \rangle - \partial_\nu \langle f \partial_\alpha, \partial_\lambda \rangle
- \partial_\nu \langle h \partial_\beta, \partial_\lambda \rangle - \langle \partial_\lambda, [(f \partial_\alpha + h \partial_\beta), \partial_\nu] \rangle - \langle \partial_\nu, [\partial_\lambda, (f \partial_\alpha + h \partial_\beta)] \rangle.
\]
By the definition of the Lie bracket and the linearity property of the metric tensor, this can be rewritten

\[ f\partial_\alpha g_{\nu\lambda} + h\partial_\beta g_{\nu\lambda} + f\partial_\lambda g_{\nu\alpha} + h\partial_\nu g_{\nu\lambda} - f\partial_\nu g_{\nu\lambda} - h\partial_\nu g_{\beta\lambda} \]

\[-\langle \partial_\lambda, ((f\partial_\alpha + h\partial_\beta))\rangle - \langle \partial_\nu, (f\partial_\alpha + h\partial_\beta)\rangle - (f\partial_\alpha + h\partial_\beta)\partial_\lambda), \]

and

\[ f\partial_\alpha g_{\nu\lambda} + h\partial_\beta g_{\nu\lambda} + f\partial_\lambda g_{\nu\alpha} + h\partial_\nu g_{\nu\lambda} - \partial_\nu f_{g\nu\lambda} - \partial_\nu h_{g\beta\lambda} - \partial_\nu f_{g\alpha\nu} + \partial_\nu h_{g\lambda\beta} - \partial_\lambda f_{g\nu\alpha} - \partial_\lambda h_{g\nu\beta}. \]

Half of these terms cancel each other out, leaving:

\[ f\partial_\alpha g_{\nu\lambda} + h\partial_\beta g_{\nu\lambda} + f\partial_\lambda g_{\nu\alpha} + h\partial_\nu g_{\nu\lambda} - \partial_\nu f_{g\nu\lambda} - \partial_\nu h_{g\beta\lambda} - \partial_\nu f_{g\alpha\nu} + \partial_\nu h_{g\lambda\beta} - \partial_\lambda f_{g\nu\alpha} - \partial_\lambda h_{g\nu\beta} = 2f \sum_\mu \Gamma^\mu_\alpha g_{\rho\nu} + 2h \sum_\rho \Gamma^\rho_\beta g_{\rho\nu}. \]

Now, we return to the Koszul formula to prove linearity in the second variable. We recall that we can without loss of generality use \( \partial_\alpha \) instead of \( X \), \( \partial_\beta + \partial_\lambda \) instead of \( Y \) and \( \partial_\nu \) instead of \( Z \). Then, equation (2.1) becomes

\[ 2\langle \nabla_{\partial_\alpha} (\partial_\beta + \partial_\lambda), \partial_\nu \rangle = \partial_\alpha \langle (\partial_\beta + \partial_\lambda), \partial_\nu \rangle + (\partial_\beta + \partial_\lambda)\langle \partial_\nu, \partial_\alpha \rangle - \partial_\nu \langle (\partial_\beta + \partial_\lambda), \partial_\alpha \rangle - (\partial_\beta + \partial_\lambda)\langle \partial_\nu, (\partial_\beta + \partial_\lambda) \rangle \]

\[-(\partial_\beta + \partial_\lambda)\langle \partial_\nu, \partial_\alpha \rangle - \langle \partial_\nu, [(\partial_\beta + \partial_\lambda), \partial_\alpha] \rangle - \langle \partial_\alpha, [\partial_\nu, (\partial_\beta + \partial_\lambda)] \rangle. \]

(2.7)

Now we prove that the right-hand side of this equation is equal to \( 2 \sum_\mu \Gamma^\mu_\alpha g_{\rho\nu} + 2 \sum_\mu \Gamma^\rho_\beta g_{\rho\nu} \). Clearly the last three terms of the right-hand side of equation (2.7) are zero. By linearity of the metric tensor, the right-hand side of (2.7) becomes

\[ \partial_\alpha g_{\beta\nu} + \partial_\alpha g_{\nu\lambda} + \partial_\beta g_{\nu\alpha} + \partial_\lambda g_{\nu\alpha} - \partial_\nu g_{\alpha\beta} - \partial_\nu g_{\nu\lambda} = 2 \sum_\mu \Gamma^\mu_\alpha g_{\rho\nu} + 2 \sum_\mu \Gamma^\rho_\alpha g_{\rho\nu}. \]

Therefore, property ii) is satisfied.

Now we prove that property iii) holds: Again using the Koszul equation (2.1), and again we remember that we can use \( \partial_\alpha \) instead of \( X \), \( f\partial_\beta \) instead of \( Y \) and \( \partial_\lambda \) instead of \( Z \). Using this, we arrive at

\[ 2\langle \nabla_{\partial_\alpha} f\partial_\beta, \partial_\lambda \rangle = \partial_\alpha \langle f\partial_\beta, \partial_\lambda \rangle + f\partial_\beta \langle \partial_\lambda, \partial_\alpha \rangle - \partial_\lambda \langle \partial_\alpha, f\partial_\beta \rangle \]

\[-(f\partial_\beta, [\partial_\lambda, \partial_\alpha]) - \langle \partial_\lambda, [f\partial_\beta, \partial_\alpha] \rangle - \langle \partial_\alpha, (f\partial_\beta, \partial_\lambda) \rangle. \]

We note that what we actually want to prove is that the right-hand side is equal to \( 2(\partial_\alpha f g_{\beta\lambda} + f \sum_\mu \Gamma^\mu_\alpha g_{\rho\nu}) \). Then we can rewrite the right-hand side of this equation.

We expand the last three terms by the definition of the Lie bracket, and by the linearity property of the metric tensor we find it to be equal to

\[ \partial_\alpha (f g_{\beta\lambda}) + f\partial_\beta g_{\alpha\lambda} - \partial_\lambda (f g_{\alpha\beta}) - \langle \partial_\lambda, (f \partial_\alpha - \partial_\alpha f) \partial_\beta \rangle - \langle \partial_\alpha, (f \partial_\lambda - \partial_\lambda f) \partial_\beta \rangle. \]
Recognizing that only one term in each Lie derivative remains nonzero, by the
symmetry and linearity properties of the metric tensor, this equation is equal
to
\[ \partial_\alpha f g_{\beta \lambda} + f \partial_\alpha g_{\beta \lambda} + f \partial_\beta g_{\alpha \lambda} - \partial_\lambda f g_{\alpha \beta} - f \partial_\lambda g_{\alpha \beta} + \partial_\alpha f g_{\lambda \beta} \]
finally, some terms cancel, leaving
\[ 2 \partial_\alpha f g_{\beta \lambda} + f \partial_\alpha g_{\beta \lambda} + f \partial_\beta g_{\alpha \lambda} - f \partial_\lambda g_{\alpha \beta} = 2(f \partial_\alpha f g_{\beta \lambda} + f \sum_\mu \Gamma^\mu_{\alpha \beta} g_{\mu \lambda}). \]
This clearly proves that the covariant derivative is Leibnizian in the second variable.

Now we prove that property iv) holds: Since \([\partial_\alpha, \partial_\beta] = 0\) for all \(\alpha, \beta\), this property simply gives us \(\Gamma^\nu_{\alpha \beta} = \Gamma^\nu_{\beta \alpha}\) in terms of the metric tensor.

Now we prove that property v) holds: It is no loss of generality to prove that
\[ \partial_\nu (\partial_\beta, \partial_\lambda) = \langle \nabla_{\partial_\nu} \partial_\beta, \partial_\lambda \rangle + \langle \partial_\beta, \nabla_{\partial_\nu} \partial_\lambda \rangle \]
or in terms of \(\Gamma\) and the metric tensor components:
\[ \partial_\nu g_{\beta \lambda} = \sum_\mu \Gamma^\mu_{\alpha \beta} g_{\mu \lambda} + \sum_\mu \Gamma^\mu_{\alpha \lambda} g_{\mu \beta}. \]
The right-hand side can be written, using equation (2.4) and the symmetry of
the metric tensor, as
\[ \sum_\mu \Gamma^\mu_{\alpha \beta} g_{\mu \lambda} + \sum_\mu \Gamma^\mu_{\alpha \lambda} g_{\mu \beta} = \frac{1}{2} (\partial_\alpha g_{\beta \lambda} + \partial_\beta g_{\alpha \lambda} - \partial_\lambda g_{\alpha \beta}) + \frac{1}{2} (\partial_\alpha g_{\lambda \beta} + \partial_\beta g_{\lambda \alpha} - \partial_\lambda g_{\alpha \beta}) = \partial_\nu g_{\beta \lambda}. \]
Thus, we have proven that the covariant derivative exists on every chart of a
manifold \(\mathcal{M}\), and that it is unique.

**Corollary 2.18** (Christoffel symbols in terms of the metric tensor). On some
chart \((\mathcal{U}, \phi = (u^1, u^2, ..., u^n))\) of a semi-Riemannian manifold \(\mathcal{M}\), the Christoffel
symbols are given in terms of the metric tensor components by:
\[ \Gamma^\nu_{\alpha \beta} = \frac{1}{2} \sum_\lambda g^{\lambda \nu} (\partial_\alpha g_{\beta \lambda} + \partial_\beta g_{\alpha \lambda} - \partial_\lambda g_{\alpha \beta}). \] (2.8)

**Proof.** This follows from the proof of Theorem 2.17, see equation (2.5).

As a last part of derivatives on semi-Riemannian manifolds, we will general-
ize the concept of the divergence to apply to tensors on a semi-Riemannian
manifold. This is defined in terms of the covariant derivative, and is therefore
related to normal divergence as the normal derivative is related to the covariant
derivative (see e.g. chapter 7 in [9]).

**Definition 2.19** (Divergence). Let \(\mathcal{M}\) be a semi-Riemannian manifold
with covariant derivative \(\nabla\). Then the divergence of a tensor \(\tau\) is defined as the
contraction
\[ \text{div}_\beta \tau = C^\beta_\alpha (\nabla_\alpha \tau). \] (2.9)
Expressed in components of the \((r, s)\) tensor \(\tau\), this becomes

\[
\nabla_{\alpha_n} \tau_{\alpha_1, \ldots, \alpha_n, \ldots, \alpha_r} = \sum_{\alpha_n} \nabla_{\beta_n} \tau_{\alpha_1, \ldots, \alpha_n, \ldots, \alpha_r}. \tag{2.10}
\]

For symmetric \((0, 2)\) and \((2, 0)\) tensors \(\tau\), we skip the subscript for the divergence, since it does not matter which divergence we take, and write the divergence of \(\tau\) as \(\text{div} \tau\).

In this paper we will mainly be interested in \(\text{div} \tau = 0\), due to the physical significance. We are chiefly interested in the case when the energy and momentum content of the universe has zero divergence, that is, when there is no energy or momentum disappearing from the universe.

### 2.3 Distance

We will talk a lot about distances, lengths of vectors and similar. Some authors dislike using the word length since a null curve has length zero by this definition. However, stretching the normal concept of length, we may as well call this \(L\) the length of a curve on a semi-Riemannian manifold.

It may be useful for most readers at this point to note that the word metric that we use regularly in this paper is not a metric in the topological sense. When we talk about the geometry of a manifold, it is the properties of the inner product on its tangent bundle we mean. This in turn gives a new meaning to integral curves on the manifold, the differential geometry of the manifold. It is in this sense that we use the words geometry and metric. In fact, at several points we will use the term ‘on the manifold’ when we in fact mean on the tangent bundle of the manifold, or perhaps even only in some local tangent space. This should not cause any confusion, and in fact what we really mean should be clear from context.

The concept of length holds up in positive-definite (metric signature with \(n = 0\)) geometry, where all nonzero vectors have nonzero length, as well as indefinite \((n \neq 0)\) geometry, where a vector may well have negative length in some sense, although of course the arc length of a curve cannot be negative.

**Definition 2.20** (Arc length). Let \(\mathcal{M}\) be some semi-Riemannian manifold, and let \(\gamma : I \to \mathcal{M}\) be a curve on some chart \((U, \phi)\). Then we define the arc length of the curve \(\gamma\) as:

\[
L = \int_I |\dot{\gamma}(s)| ds = \int_I |\langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle|^{1/2} ds,
\]

which with our notation for the metric tensor, this can be rewritten:

\[
L = \int_I \left| \sum_{\alpha\beta} g_{\alpha\beta} \frac{du^\alpha \circ \gamma(s)}{ds} \frac{du^\beta \circ \gamma(s)}{ds} \right|^{1/2} ds.
\]

This arc length of a curve is unchanged by monotone reparametrization (see, e.g. chapter 5 in [16]).

Since we now have the concepts of both lengths of curves and length of vectors on a semi-Riemannian manifold \(\mathcal{M}\), we have the tools necessary to talk
about minimizing the arc length $L$ between two points on $M$. In normal Euclidean space, the shortest distance between two points is a straight line.

We will generalize this concept of a straight line in an intuitive fashion, and so we will be able to visualize even high-dimensional, general manifolds as analogous to a curved Euclidean space.

### 2.4 Geodesics and parallel transport

On some semi-Riemannian manifold $M$, the concept of a shortest path or straight line can be generalized. It can be thought of as a curve $\gamma$, whose vector field $\dot{\gamma}$ is parallel. This concept of a straight line, or geodesic, is widely used in physical applications.

Let the tangent vector field along the curve $\gamma$ be denoted by $\dot{\gamma}$. This $\dot{\gamma}$ is therefore the directional derivative. Let us parametrize the curve using an affine parameter $s$, so that, using this coordinate, $\gamma = \gamma(s)$. Then the directional derivative, the vector field $\dot{\gamma}$, is the derivative, $d\gamma/ds$. For $\dot{\gamma}$ to be parallel, we will require that $\nabla \dot{\gamma} \dot{\gamma} = 0$.

**Definition 2.21** (Geodesic). Let $M$ be a semi-Riemannian manifold and let $\gamma: [p, q] \rightarrow M$ be a curve on $M$. Then $\gamma$ is a geodesic if $\nabla \dot{\gamma} \dot{\gamma} = 0$.

We shall now formulate the differential equations that describe these geodesics, and further we will prove the uniqueness and existence of geodesics.

**Theorem 2.22** (Existence and uniqueness of geodesic equations). Let $M$ be a semi-Riemannian manifold and let $\gamma: [p, q] \rightarrow M$ be a curve on $M$. If $\gamma$ is a geodesic the coordinate functions $u^\alpha \circ \gamma$ on some chart $(U, \phi)$ satisfy the geodesic equations:

$$\frac{d^2(u^\lambda \circ \gamma(s))}{ds^2} + \sum_\alpha \sum_\beta \Gamma^\lambda_{\alpha \beta} \frac{d(u^\alpha \circ \gamma(s))}{ds} \frac{d(u^\beta \circ \gamma(s))}{ds} = 0. \quad (2.11)$$

For each geodesic, these equations have unique solutions.

**Proof.** First, we will prove the equivalence between the definition of a geodesic, that is $\nabla \dot{\gamma} \dot{\gamma} = 0$, and the geodesic equations. Then we will argue for the uniqueness and existence of the solutions to the geodesic equations.

**First, we prove equivalence:** For a curve $\gamma$, the vector field along the curve, $\dot{\gamma}$, can be rewritten

$$\dot{\gamma} = \sum_\alpha \dot{\gamma}^\alpha \partial_\alpha.$$

Using this with the definition of a geodesic gives us from the Leibnizian property of the covariant derivative:

$$\nabla \dot{\gamma} = \nabla \gamma \sum_\alpha \dot{\gamma}^\alpha \partial_\alpha = \sum_\alpha (\nabla_\gamma \dot{\gamma}^\alpha) \partial_\alpha + \sum_\alpha \dot{\gamma}^\alpha (\nabla_\gamma \partial_\alpha).$$

Now, for the first sum, it is clear that the covariant derivative relative to the vector curve $\dot{\gamma}$ is simply the derivative of the vector field on the curve relative to the parameter $s$, i.e. $\nabla_\gamma \dot{\gamma}^\alpha = \ddot{\gamma}^\alpha$. 

19
Using this along with the linearity property of $\nabla_\gamma = \nabla_{\Sigma^\beta \partial_\beta}$, we can rewrite this equation as

$$\nabla_\gamma \dot{\gamma} = \sum_\alpha \dot{\gamma}^\alpha \partial_\alpha + \sum_\alpha \dot{\gamma}^\alpha (\nabla_\Sigma^\beta \partial_\beta \partial_\alpha),$$

and using property ii) of the definition of the covariant derivative, this equation becomes

$$\nabla_\gamma \dot{\gamma} = \sum_\alpha \dot{\gamma}^\alpha \partial_\alpha + \sum_\alpha \sum_\beta \dot{\gamma}^\alpha \dot{\gamma}^\beta (\nabla_\partial_\beta \partial_\alpha).$$

Now, using Definition 2.21 on the left-hand side of this equation, and Definition 2.15, we obtain

$$0 = \sum_\alpha \dot{\gamma}^\alpha \partial_\alpha + \sum_\alpha \sum_\beta \dot{\gamma}^\alpha \dot{\gamma}^\beta \sum_\lambda \Gamma^\lambda_{\beta\alpha} \partial_\lambda.$$

and if we change the index in the first term from $\alpha$ to $\lambda$, we get

$$0 = \sum_\lambda \left( \dot{\gamma}^\lambda + \sum_\alpha \sum_\beta \dot{\gamma}^\alpha \dot{\gamma}^\beta \Gamma^\lambda_{\beta\alpha} \right) \partial_\lambda.$$

Clearly equivalence is only held in general if each term of this sum over $\lambda$ is zero, when the components of this vector are all zero. That is,

$$\dot{\gamma}^\lambda + \sum_{\alpha,\beta} \dot{\gamma}^\alpha \dot{\gamma}^\beta \Gamma^\lambda_{\beta\alpha} = 0.$$

Since $\gamma^\alpha = u^\alpha \circ \gamma$, these are the geodesic equations and we have derived our geodesic equations from the definition of the geodesic.

**Proving existence and uniqueness of the solutions:** To prove uniqueness and existence, we need only note that putting $y = \dot{\gamma}$, we get the system of first-order equations

$$\begin{align*}
y^\lambda &= \dot{\gamma}^\lambda, \\
y^\lambda &= -\sum_\alpha \sum_\beta y^\alpha y^\beta \Gamma^\lambda_{\beta\alpha},
\end{align*}$$

and by the existence and uniqueness theorem for first-order ordinary differential equations, there exists a unique solution for this system of equations, for every given set of initial conditions $\gamma(s_0) = p, \dot{\gamma}(s_0) = X_p$ imposed upon it (see, e.g. Theorem 4.1 in [20]).

This next example is an immediate application of equation (2.11). In it, we study the geodesic equations for an orthogonal coordinate system, or more accurately, an orthogonal coordinate basis.

Expressed in terms of the metric tensor components, orthogonality requires that $g_{\alpha\beta} = 0$ for $\alpha \neq \beta$. This leads to the geodesic equations, which are normally quite complicated, to take a much simpler form.
Example 2.23 (Geodesic equations in orthogonal coordinates). In this example, inspired by Exercise 3.15 in [16], we will examine the geodesic equations in orthogonal coordinates closely. Orthogonal coordinates are in practice very common, for example the Schwarzschild solution to the Einstein field equations is orthogonal (see (5.2)). Similarly, both examples in section 4.3 are orthogonal.

Let \( \mathcal{M} \) be a semi-Riemannian manifold with metric tensor \( g \). For a chart \((U, \phi)\), the coordinate basis is said to be orthogonal if the metric tensor relative to \( \phi \) has the property \( g_{\alpha\beta} = 0 \) if \( \alpha \neq \beta \).

Notice that this makes the sum in the definition of the Christoffel symbols from (2.8), vanish. Equation (2.8) then becomes

\[
\Gamma^\lambda_{\alpha\beta} = \frac{1}{2} g^{\lambda\lambda} (\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\lambda\alpha} - \partial_{\lambda} g_{\alpha\beta})
\]

This allows us to make the general geodesic equations, equation (2.11), a little more manageable as

\[
\frac{d^2 u^\lambda}{ds^2} + \frac{1}{2} \sum_\alpha \sum_\beta g^{\lambda\lambda} \partial_\alpha g_{\beta\lambda} \frac{du^\alpha \circ \gamma(s)}{ds} \frac{du^\beta \circ \gamma(s)}{ds} \\
\quad + \frac{1}{2} \sum_\alpha \sum_\beta g^{\lambda\lambda} \partial_\beta g_{\lambda\alpha} \frac{du^\alpha \circ \gamma(s)}{ds} \frac{du^\beta \circ \gamma(s)}{ds} \\
- \frac{1}{2} \sum_\alpha \sum_\beta g^{\lambda\lambda} \partial_\lambda g_{\alpha\beta} \frac{du^\alpha \circ \gamma(s)}{ds} \frac{du^\beta \circ \gamma(s)}{ds} = 0.
\]

From the assumption of orthogonality, we have that in the second term, \( \beta = \lambda \) is the only nonzero case, in the third only \( \alpha = \lambda \) is nonzero and in the fourth, only \( \alpha = \beta \) is nonzero.

Replacing the index \( \alpha \) for \( \beta \) in the third term, we see that the second term and third term are equal and we can rewrite these equations as

\[
\frac{d^2 u^\lambda \circ \gamma(s)}{ds^2} + g^{\lambda\lambda} \sum_\alpha \partial_\alpha g_{\beta\lambda} \frac{du^\alpha \circ \gamma(s)}{ds} \frac{du^\beta \circ \gamma(s)}{ds} \\
- \frac{1}{2} g^{\lambda\lambda} \sum_\alpha \partial_\lambda g_{\alpha\beta} \frac{du^\alpha \circ \gamma(s)}{ds} \frac{du^\alpha \circ \gamma(s)}{ds} = 0.
\]

Lastly, we multiply by \( g_{\lambda\lambda} \) and by some straightforward calculations we obtain the geodesic equations for orthogonal coordinates

\[
\frac{d}{ds} \left( g_{\lambda\lambda} \frac{du^\lambda}{ds} \right) - \frac{1}{2} \sum_\alpha \partial_\lambda g_{\alpha\alpha} \left( \frac{du^\alpha}{ds} \right)^2 = 0. \tag{2.12}
\]

These equations are much simpler to use and therefore very practical for all practical cases we will consider in this paper. In fact, all metric tensors presented in this paper are orthogonal, and therefore this version of the geodesic equations hold on all manifolds considered.

Next, we present a proposition for geodesic parametrization. We will not need to use the results explicitly, but it will provide insight in how we are allowed to parametrize geodesics.
This will be particularly useful for the reader in the following chapters, where we will consider geodesics parametrized by what is called proper time (see e.g. chapter 3 and 4 in [4] for more details on the physical significance of this proposition).

**Proposition 2.24** (Geodesic parametrization). Let $\gamma: I \to \mathcal{M}$ be a nonconstant geodesic on some semi-Riemannian manifold $\mathcal{M}$. A reparametrization $\tilde{\gamma}$ of $\gamma$ using a function $f: I \to J$ is a geodesic if $f(s) = as + b, a,b \in \mathbb{R}$.

**Proof.** If $\gamma = \gamma(s)$ is a geodesic and $f$ is some real-valued function, then the reparametrized curve $\tilde{\gamma} = \gamma \circ f$ needs to satisfy the geodesic equations, and $\nabla_{\dot{\tilde{\gamma}}}\dot{\tilde{\gamma}} = 0$ becomes, writing $d\tilde{\gamma}/ds$ as $\dot{\tilde{\gamma}}$:

$$\nabla_{\dot{\tilde{\gamma}}}\dot{\tilde{\gamma}} = \nabla_{\dot{\gamma}}\left(\frac{d\gamma}{df}\frac{df}{ds}\right) = \frac{df}{ds}\left(\nabla_{\dot{\gamma}}\frac{d\gamma}{df}\right) + \frac{d\gamma}{df}\left(\nabla_{\dot{\gamma}}\frac{df}{ds}\right).$$

Since obviously $\nabla_{\dot{\gamma}}(d\gamma/df) = 0$, it is required that $\nabla_{\dot{\gamma}}(df/df) = d^2f/df^2 = 0$, and that requires from the function $f$ that $f = as + b$. \qed

### 2.5 Curvature

Now, having defined the length of a curve and a geodesic, we can further understand a semi-Riemannian manifold if we think of the geometry as defined by how we measure distance between points. However, there is another useful tool when we consider the geometry of a manifold, and that is the curvature of the manifold.

So what is the purpose of the curvature tensor? Well, visualizing the parallel transport of a vector around a closed curve $\gamma$, one would expect the change in the vector to be zero, and in Euclidean space this is true. However, this is not true in general, and generally one may expect a difference which depends on the curvature. Therefore, we can think of the curvature as somewhat analogous to the difference between the start vector and end vector when parallel transported in a closed loop on the manifold.

In fact, we may define flatness at a point $p$ of a semi-Riemannian manifold $\mathcal{M}$ by every component of the curvature tensor being equal to zero at that point (see e.g. chapter 13 in [7] for more details).

**Definition 2.25** (Curvature tensor). The curvature tensor (sometimes called the Riemann curvature tensor), is a $(1,3)$ tensor field, $R: \mathfrak{X}(\mathcal{M})^3 \to \mathfrak{X}(\mathcal{M})$ such that

$$R_{XY}Z = \nabla_{[X,Y]}Z - [\nabla_X, \nabla_Y]Z,$$

where $X,Y,Z \in \mathfrak{X}(\mathcal{M})$.

**Proposition 2.26** (Curvature tensor properties). The curvature tensor $R$ has the following properties

i. Skew-symmetry: $R_{XY} = -R_{YX}$,

ii. First Bianchi identity: $R_{XY}Z + R_{YZ}X + R_{ZX}Y = 0$,

iii. Metric skew-symmetry: $(R_{XY}Z, W) = -(R_{XY}W, Z)$.  

22
for all $X, Y, Z, W \in \mathcal{X}(\mathfrak{M})$

**Proof.** Proving property $i$): The property $i$) is obvious since the Lie bracket is skew-symmetric, and thus from the definition of the curvature tensor and the covariant derivative, with very simple algebra this property follows.

Proving property $ii$): The second property is perhaps not so obvious, but rewriting the left side of this equality by the definition of the curvature tensor as

$$
\begin{align*}
R_{XY}Z + R_{YZ}X + R_{ZX}Y &= \nabla_{[X,Y]}Z - (\nabla_X \nabla_Y - \nabla_Y \nabla_X)Z \\
&+ \nabla_{[Y,Z]}X - (\nabla_Y \nabla_Z - \nabla_Z \nabla_Y)X + \nabla_{[Z,X]}Y - (\nabla_Z \nabla_X - \nabla_X \nabla_Z)Y \\
&= \nabla_{[X,Y]}Z + \nabla_{[Y,Z]}X + \nabla_{[Z,X]}Y - \nabla_X (\nabla_Y Z - \nabla_Z Y) \\
&- \nabla_Y (\nabla_Z X - \nabla_X Z) - \nabla_X (\nabla_Y Y - \nabla_Y X).
\end{align*}
$$

From the torsion-free property of the covariant derivative and the definition of the Lie bracket we get that

$$
\nabla_{[X,Y]}Z + \nabla_{[Y,Z]}X + \nabla_{[Z,X]}Y - \nabla_X [Y, Z] - \nabla_Y [Z, X] - \nabla_Z [X, Y] = [[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0,
$$

by the Jacobi identity (see e.g. page 92 in [7]).

Proving property $iii$): Now, it is sufficient to prove that

$$
\langle R_{XY} \Lambda, \Lambda \rangle = 0 \quad \text{for all} \quad \Lambda,
$$

since then, for $\Lambda = X + Y$,

$$
\langle R_{XY} (X + Y, X + Y) \rangle = \langle R_{XY} X, X \rangle + \langle R_{XY} Y, Y \rangle + \langle R_{XY} X, Y \rangle + \langle R_{XY} Y, X \rangle = \langle R_{XY} X, Y \rangle + \langle R_{XY} Y, X \rangle = 0.
$$

From the definition of the curvature tensor and the linearity of the metric tensor we have that

$$
\langle R_{XY} \Lambda, \Lambda \rangle = \langle \nabla_{[X,Y]} \Lambda, \Lambda \rangle - \langle \nabla_X \nabla_Y \Lambda, \Lambda \rangle + \langle \nabla_Y \nabla_X \Lambda, \Lambda \rangle,
$$

and from the metric compatibility property of Definition 2.14, we get for each term of the right hand side of this equation

- $\langle \nabla_{[X,Y]} \Lambda, \Lambda \rangle = [X, Y] \langle \Lambda, \Lambda \rangle / 2$,

  $$
  \langle \nabla_X \nabla_Y \Lambda, \Lambda \rangle = X \langle \nabla_Y \Lambda, \Lambda \rangle - \langle \nabla_X \Lambda, \nabla_Y \Lambda \rangle = XY \langle \Lambda, \Lambda \rangle / 2 - \langle \nabla_X \Lambda, \nabla_Y \Lambda \rangle,
  $$

  $$
  \langle \nabla_Y \nabla_X \Lambda, \Lambda \rangle = Y \langle \nabla_X \Lambda, \Lambda \rangle - \langle \nabla_Y \Lambda, \nabla_X \Lambda \rangle = YX \langle \Lambda, \Lambda \rangle / 2 - \langle \nabla_X \Lambda, \nabla_Y \Lambda \rangle.
  $$

Hence,

$$
\langle R_{XY} \Lambda, \Lambda \rangle = [X, Y] \langle \Lambda, \Lambda \rangle / 2 - XY \langle \Lambda, \Lambda \rangle / 2 + YX \langle \Lambda, \Lambda \rangle / 2 = 0,
$$

Thus, $\langle R_{XY} \Lambda, \Lambda \rangle = 0$. \qed

Mostly in calculations, we are interested in differential equations of real-valued functions. Therefore we are interested in the real-valued components of the curvature tensor, which we will use in later chapters and examples.
Theorem 2.27 (Curvature tensor components). For some chart \((U, \phi)\) on a semi-Riemannian manifold \(\mathcal{M}\), the components of \(\mathbf{R}\) relative to the coordinate system are given by

\[
\mathbf{R}_{\partial_\alpha, \partial_\beta} \partial_\lambda = \sum_\mu R^\mu_{\lambda\alpha\beta} \partial_\mu,
\]

where the real-valued functions \(R^\mu_{\lambda\alpha\beta}\) are given by

\[
R^\mu_{\lambda\alpha\beta} = \partial_\alpha \Gamma^\mu_{\beta\lambda} - \partial_\beta \Gamma^\mu_{\alpha\lambda} + \sum_\nu \Gamma^\mu_{\alpha\nu} \Gamma^\nu_{\beta\lambda} - \sum_\nu \Gamma^\mu_{\beta\nu} \Gamma^\nu_{\alpha\lambda}.
\]

Proof. In terms of coordinate vector field \((\partial_1, \partial_2, ..., \partial_n)\), the definition of the curvature tensor becomes

\[
\mathbf{R}_{\partial_\alpha, \partial_\beta} \partial_\lambda = \nabla_{[\partial_\alpha, \partial_\beta]} \partial_\lambda - [\nabla_{\partial_\alpha}, \nabla_{\partial_\beta}] \partial_\lambda,
\]

and from the definition of the Christoffel symbols, and since \([\partial_\alpha, \partial_\beta] = 0\) for all \(\alpha, \beta\), we get

\[
\mathbf{R}_{\partial_\alpha, \partial_\beta} \partial_\lambda = \nabla_{\partial_\alpha} \nabla_{\partial_\beta} \partial_\lambda - \nabla_{\partial_\beta} \nabla_{\partial_\alpha} \partial_\lambda
\]

\[
= \nabla_{\partial_\alpha} \sum_\mu \Gamma^\mu_{\alpha\lambda} \partial_\mu - \nabla_{\partial_\beta} \sum_\mu \Gamma^\mu_{\beta\lambda} \partial_\mu.
\]

Since \(\nabla_{\partial_\alpha} f = \partial_\alpha f\) for real-valued functions \(f\), we have from the Leibniz property of the covariant derivative

\[
\mathbf{R}_{\partial_\alpha, \partial_\beta} \partial_\lambda = \nabla_{\partial_\alpha} \sum_\mu \Gamma^\mu_{\alpha\lambda} \partial_\mu - \nabla_{\partial_\beta} \sum_\mu \Gamma^\mu_{\beta\lambda} \partial_\mu = \sum_\mu \left( \partial_\beta \Gamma^\mu_{\alpha\lambda} \partial_\mu + \Gamma^\mu_{\alpha\lambda} \nabla_{\partial_\beta} \partial_\mu \right)
\]

\[
- \sum_\mu \left( \partial_\alpha \Gamma^\mu_{\beta\lambda} \partial_\mu + \Gamma^\mu_{\beta\lambda} \nabla_{\partial_\alpha} \partial_\mu \right)
\]

and expanding,

\[
\mathbf{R}_{\partial_\alpha, \partial_\beta} \partial_\lambda = \sum_\mu \left( \partial_\beta \Gamma^\mu_{\alpha\lambda} \partial_\mu + \Gamma^\mu_{\alpha\lambda} \sum_\nu \Gamma^\nu_{\beta\mu} \partial_\nu \right)
\]

\[
- \sum_\mu \left( \partial_\alpha \Gamma^\mu_{\beta\lambda} \partial_\mu + \Gamma^\mu_{\beta\lambda} \sum_\nu \Gamma^\nu_{\alpha\mu} \partial_\nu \right).
\]

Rearranging by changing the place of the outer and inner sums and relabeling \(\mu\) to \(\nu\) in the first and third terms,

\[
\mathbf{R}_{\partial_\alpha, \partial_\beta} \partial_\lambda = \sum_\nu \left( \partial_\beta \Gamma^\nu_{\alpha\lambda} + \sum_\mu \Gamma^\mu_{\alpha\lambda} \Gamma^\nu_{\beta\mu} - \partial_\alpha \Gamma^\nu_{\beta\lambda} - \sum_\mu \Gamma^\mu_{\alpha\lambda} \Gamma^\nu_{\beta\mu} \right) \partial_\nu.
\]

From this we identify the components of the curvature tensor, and see that equation (2.13) is satisfied, with the real-valued component functions given by

\[
R^\mu_{\lambda\alpha\beta} = \partial_\alpha \Gamma^\mu_{\beta\lambda} - \partial_\beta \Gamma^\mu_{\alpha\lambda} + \sum_\nu \Gamma^\mu_{\alpha\nu} \Gamma^\nu_{\beta\lambda} - \sum_\nu \Gamma^\mu_{\beta\nu} \Gamma^\nu_{\alpha\lambda}.
\]

\[\square\]
We will now introduce a last property of the curvature tensor, usually called the second Bianchi identity.

**Theorem 2.28** (Second Bianchi identity). Let \( \mathfrak{M} \) be a semi-Riemannian manifold with covariant derivative \( \nabla \) and let \( X, Y, Z \in \mathfrak{X}(\mathfrak{M}) \). Then:

\[
\nabla_X R_{Y,Z} + \nabla_Z R_{X,Y} + \nabla_Y R_{Z,X} = 0,
\]

(2.14)

and in component form

\[
\nabla_\partial_\alpha R^\alpha_{\beta\mu\nu} + \nabla_\partial_\beta R^\beta_{\alpha\mu\nu} + \nabla_\partial_\mu R^\mu_{\beta\alpha\nu} = 0.
\]

(2.15)

**Proof.** Taking inspiration from the proof of e.g. Theorem 13.20 in [7] and Proposition 37 in [16], we look at \( \nabla_X(\nabla_Y Z) \), by the Leibnizian property of the covariant derivative:

\[
\nabla_X(\nabla_Y Z) = (\nabla_X \nabla_Y Z) W + R_{Y,Z}(\nabla_X W).
\]

Rearranging this equation, we get

\[
(\nabla_X \nabla_Y Z) W = \nabla_X(\nabla_Y Z) W - R_{Y,Z}(\nabla_X W).
\]

Then in terms of the Lie bracket, this becomes

\[
\]

and doing the same for the other two terms of the initial equation (2.14), we get

\[
(\nabla_Z \nabla_X Y) W = [\nabla_Z, R_{X,Y}] W = [\nabla_Z, \nabla_X Y] - [\nabla_Y, \nabla_X] W,
\]

and

\[
(\nabla_Y \nabla_Z X) W = [\nabla_Y, R_{Z,X}] W = [\nabla_Y, \nabla_Z X] - [\nabla_Z, \nabla_X] W.
\]

Now we add these three equations together, and get

\[
\]

By linearity of the Lie bracket, we find that this is equal to

\[
(\nabla_X \nabla_Y Z) W + (\nabla_Z \nabla_X Y) W + (\nabla_Y \nabla_Z X) W
\]

\[
= [\nabla_X, [\nabla_Y, \nabla_Z]] W + [\nabla_Z, [\nabla_X, \nabla_Y]] W + [\nabla_Y, [\nabla_Z, \nabla_X]] W.
\]

With no loss of generality, as in previous proofs, we assume that the Lie bracket of any vectors \([X, Y]\) is zero, as with a coordinate basis. Then the last equation becomes

\[
(\nabla_X \nabla_Y Z) W + (\nabla_Z \nabla_X Y) W + (\nabla_Y \nabla_Z X) W
\]

\[
= -[\nabla_X, [\nabla_Y, \nabla_Z]] W - [\nabla_Z, [\nabla_X, \nabla_Y]] W - [\nabla_Y, [\nabla_Z, \nabla_X]] W.
\]

This is zero by the Jacobi identity, and therefore Theorem 2.28 holds (see e.g. page 92 in [7]).
Now, we will introduce two more quantities derived from the curvature tensor, also related to the curvature and therefore the geometry of a semi-Riemannian manifold (see e.g. chapter 7 in [9] for a lengthier discussion of the Ricci and scalar curvatures).

**Definition 2.29 (Ricci tensor).** The **Ricci tensor**, \( \text{Ric} \), on a semi-Riemannian manifold \( \mathcal{M} \) with curvature tensor \( R \), is defined to be the contraction \( C^1_3(R) \in \mathfrak{T}^0_2(\mathcal{M}) \). Its components are given by

\[
\text{Ric}_{\alpha\beta} = \sum_{\lambda} R^\lambda_{\alpha\beta\lambda}.
\]

For use in Theorem 4.4, and for mathematical understanding of the Ricci tensor, the following lemma is useful.

**Lemma 2.30 (Ricci tensor symmetry).** The Ricci tensor \( \text{Ric} \) is symmetric, i.e. \( \text{Ric}_{\alpha\beta} = \text{Ric}_{\beta\alpha} \).

**Proof.** In terms of the components of the curvature tensor, the first Bianchi identity (property ii) in Proposition 2.26) becomes

\[
R^\mu_{\alpha\beta\lambda} + R^\mu_{\beta\lambda\alpha} + R^\mu_{\lambda\alpha\beta} = 0. \tag{2.16}
\]

The skew-symmetry property (property i) in Proposition 2.26) is, in terms of the components of the curvature tensor

\[
R^\mu_{\alpha\beta\lambda} = -R^\mu_{\alpha\lambda\beta}.
\]

Now contracting equation (2.16) gives us

\[
\sum_{\mu} \left( R^\mu_{\alpha\beta\mu} + R^\mu_{\beta\mu\alpha} + R^\mu_{\mu\alpha\beta} \right) = \text{Ric}_{\alpha\beta} - \text{Ric}_{\beta\alpha} + \sum_{\mu} R^\mu_{\mu\alpha\beta} = 0. \tag{2.17}
\]

This last term can be written

\[
\sum_{\mu} R^\mu_{\mu\alpha\beta} = \sum_{\mu} \left( \partial_\beta \Gamma^\mu_{\alpha\mu} - \partial_\alpha \Gamma^\mu_{\beta\mu} + \sum_\nu \Gamma^\mu_{\beta\nu} \Gamma^\nu_{\alpha\mu} - \sum_\nu \Gamma^\mu_{\alpha\nu} \Gamma^\nu_{\beta\mu} \right),
\]

or equivalently,

\[
\sum_{\mu} R^\mu_{\mu\alpha\beta} = \sum_{\mu} \left( \partial_\beta \Gamma^\mu_{\alpha\mu} - \partial_\alpha \Gamma^\mu_{\beta\mu} \right) + \sum_{\mu} \sum_\nu \Gamma^\mu_{\beta\nu} \Gamma^\nu_{\alpha\mu} - \sum_{\mu} \sum_\nu \Gamma^\mu_{\alpha\nu} \Gamma^\nu_{\beta\mu}.
\]

These last two terms are clearly equal, and by equation (2.8),

\[
\sum_{\mu} R^\mu_{\mu\alpha\beta} = \frac{1}{2} \partial_\beta \left( \sum_{\mu} \sum_\lambda g^{\lambda\mu} \left( \partial_\alpha g_{\mu\lambda} + \partial_\mu g_{\lambda\alpha} - \partial_\lambda g_{\alpha\mu} \right) \right)
\]

\[
- \frac{1}{2} \left( \partial_\alpha \sum_{\mu} \sum_\lambda g^{\lambda\mu} \left( \partial_\beta g_{\mu\lambda} + \partial_\mu g_{\lambda\beta} - \partial_\lambda g_{\beta\mu} \right) \right).
\]
It is clear that only two terms remain of the six terms in the parentheses, giving us
\[
\sum_{\mu} R^\mu_{\mu\alpha\beta} = \frac{1}{2} \partial_\beta \left( \sum_{\mu} \sum_{\lambda} g^{\lambda\mu} \partial_\alpha g_{\mu\lambda} \right) - \frac{1}{2} \partial_\alpha \left( \sum_{\mu} \sum_{\lambda} g^{\lambda\mu} \partial_\beta g_{\mu\lambda} \right), \quad (2.18)
\]
By Theorem 7.3.3 in [9], we have
\[
\sum_{\mu} \sum_{\lambda} g^{\lambda\mu} \partial_\alpha g_{\mu\lambda} = \frac{\partial_\alpha \det [g_{\alpha\mu}]}{\det [g_{\alpha\mu}]} = \partial_\alpha \log |\det [g_{\alpha\mu}]|,
\]
where \(\det [g_{\alpha\mu}]\) is the determinant of \(g_{\alpha\mu}\). Then we can rewrite equation (2.18) as
\[
\sum_{\mu} R^\mu_{\mu\alpha\beta} = \frac{1}{2} \partial_\beta \partial_\alpha \log |\det [g_{\alpha\mu}]| - \frac{1}{2} \partial_\alpha \partial_\beta \log |\det [g_{\beta\mu}]|.
\]
Since this is zero, equation (2.17) becomes
\[
\text{Ric}_{\alpha\beta} - \text{Ric}_{\beta\alpha} = 0.
\]
This proves that the Ricci tensor is symmetric. □

Now we introduce the Ricci scalar or scalar curvature. It is a measure of the curvature of a semi-Riemannian manifold. At each point on the manifold, it gives a scalar determined by the geometry of the semi-Riemannian manifold.

**Definition 2.31 (Ricci scalar).** The *Ricci scalar* (sometimes called scalar curvature) \(S\), on a semi-Riemannian manifold \(M\) with curvature tensor \(R\), is defined to be the contraction of the tensor \(\uparrow^1 \text{Ric}\), i.e. \(C^1 (\uparrow^1 \text{Ric}) \in C^\infty (M)\).
Now we have studied the properties of semi-Riemannian manifolds, and we have treated curvature, distance and geodesics. So, we are ready to deal with some specific types of semi-Riemannian manifolds.

First, we will consider Riemannian manifolds which are a type of semi-Riemannian manifold that are more commonly treated than others. This family of manifolds have a positive definite metric tensor, which makes them similar to the Euclidean space.

Second, we will treat Lorentzian manifolds, which are a kind of semi-Riemannian manifold with plenty of physical applications, and this is the main reason that we are interested in them. They are used as a model for the physical spacetime in the theories of relativity.

A Venn diagram for understanding the differences and relationships between these different groups of manifolds is presented in figure 3.1.

We will end this chapter with two examples of these manifolds, along with calculations of the curvature tensor and some geodesics.

3.1 Riemannian geometry

Definition 3.1 (Riemannian manifold). A Riemannian manifold is a semi-Riemannian manifold of dimension $n$ with metric signature $(n,0)$.

From this definition, we call the metric tensor on a Riemannian manifold positive definite. Examples of Riemannian manifolds are common, the most notable one perhaps being Euclidean space, $\mathbb{R}^n$ equipped with the dot product, as described in Example 2.8.

Riemannian geometry is commonly used without taking the path through the more general semi-Riemannian geometry (see e.g. [8]). Instead many authors choose to talk about semi-Riemannian as an extension of the Riemannian case, with a more relaxed definition of the metric tensor.

As mentioned before, we see the Riemannian manifold as a special case, where the weaker requirement of nondegeneracy of the metric tensor is typically replaced by a requirement for the metric tensor to be positive definite, i.e.

$$g(X,Y) > 0 \text{ for all } X,Y \neq 0 \in \mathfrak{X}(\mathbb{M}).$$
3.2 Lorentzian geometry

We will now consider Lorentzian manifolds, a type of semi-Riemannian manifold that is used extensively to model the physical spacetime of the universe. In this section, we give a concise summary of the way we will look at vectors and curves in Lorentzian manifolds. The reader of a mathematical persuasion may think of lightlike, causal future and so on as merely mathematical label. The physical mathematician may think of these families of vectors and curves as related to physical reality.

**Definition 3.2 (Lorentzian manifold).** A Lorentzian manifold is a semi-Riemannian manifold of dimension $n$ with metric signature $(1, n-1)$.

**Definition 3.3 (Causal character of tangent vectors).** On a Lorentzian manifold $\mathcal{M}$ with metric tensor $g$, a tangent vector $X_p$ is spacelike, null or timelike, characterized accordingly:

i. **Timelike:**

$$\langle X_p, X_p \rangle > 0 \text{ while } X_p \neq 0,$$

otherwise $\langle X_p, X_p \rangle = 0$.

ii. **Null or lightlike:**

$$\langle X_p, X_p \rangle = 0 \text{ and } X_p \neq 0,$$

iii. **Spacelike:**

$$\langle X_p, X_p \rangle < 0.$$

This could be called the way we measure the length of a vector on $\mathcal{M}$. Perhaps we need to extend our view of what length is, since a nonzero vector can have zero length. This is significant both mathematically and physically. Mathematically, by Nash’s embedding theorem, any $N$-dimensional Riemannian manifold can be isometrically embedded into $\mathbb{R}^n$ ([14]). However, this is definitely not true for Lorentzian manifolds.

What we mean by isometrical embedding here is that there exists a function on every $N$-dimensional Riemannian manifold $\mathcal{M}$, $f: \mathcal{M} \to \mathbb{R}^n$ such that $g(X, Y) = g(df(X), df(Y))$. This describes an embedding of the metric of a Riemannian manifold. In this it differs from (but is based on) the Whitney embedding theorem, which describes the embedding of a manifold and its tangent bundle ([22]).

Remember the definition of a semi-Euclidean space $\mathbb{R}^n_{\nu}$ (Example 2.9) and that in particular, $\mathbb{R}^1_{\nu}$ is called a Minkowski space. Note that many other authors vary this notation by denoting a semi-Euclidean space by (in our system) $\mathbb{R}^{n+\nu}$ or sometimes the sign convention is reversed so that it is written $\mathbb{R}^{\nu}_{(n+\nu)}$.

Now, what is interesting about the mathematical perspective of an $N$-dimensional Lorentzian manifold is that it cannot in general be embedded in a $\mathbb{R}^1_{\nu}$ manifold, sometimes called $L^n$ ([13]). However, any $N$-dimensional semi-Riemannian manifold can be isometrically embedded into the semi-Euclidean space $\mathbb{R}^n_{\nu}$, of sufficiently large $n$ and $\nu$ ([13]). Lorentzian manifolds are also interesting from a physical perspective due to the applications of Lorentzian manifolds to relativity. It is from this perspective that the previous definition of the causal character of vectors becomes interesting.
Remark (Causal character of curves). Curves, or from a physical perspective the paths traced by particles, are called timelike curves, null curves and spacelike curves if the tangent vector at all points in the curve is timelike, null or spacelike, respectively. Similarly, a coordinate function is called a timelike coordinate if its coordinate curve is timelike at all points. It should perhaps be obvious that all Lorentzian manifolds have only one timelike coordinate, usually denoted by $t$. Causal curves (or future-directed causal curves) are curves with a tangent vector that at all points is either null or timelike. They have the property that at all points on the curve, the tangent vector is future-directed, i.e. it points towards positive time. These causal curves are interesting because they denote physically possible chains of events. That is, normally only causal curves are allowed in physical spacetime.

We may think of a (future-directed) causal curve as a curve from points in the past to points in the future. A spacelike curve, on the other hand, is a curve connecting points that are not in each others’ past or future.

This means that a causal curve describes the path of a particle: a timelike curve describes the path of a massive particle and a null curve describes the path of a massless particle.

We should now define what we mean by past and future. These concepts are intuitively easy to understand, but we should be careful when talking about past and future on a manifold (or on the tangent bundle of a manifold) (see e.g. [11]).

Definition 3.4 (Future and past and elsewhere). The future of a point $p$ is the set of all points $q$ such that there is a causal curve $\gamma: [p,q] \to \mathbb{R}$.

The past of a point $p$ is the set of all points $q$ such that there is a causal curve $\gamma: [q,p] \to \mathbb{R}$.

The elsewhere of a point $p$ is the set of all points $q$ such that there is no causal curves $\gamma: [q,p] \to \mathbb{R}$ or $\gamma: [p,q] \to \mathbb{R}$.

1Commonly when talking about physical spacetime, a point is called an event, since the point includes a timelike coordinate.
3.2 Hyperspace of the present

Figure 3.2: The lightcone for a point \( p \).

Usually these concepts are visualized as a cone in \( \mathbb{R}^3 \), letting the coordinates \( x, y \) and \( t \) be two spacelike and one timelike coordinate (see figure 3.2). For each point \( p \), the lightcone (or nullcone) of \( p \) has as a boundary all points \( q \) with lightlike curves from \( p \) to \( q \) (for the future) and \( q \) to \( p \) (for the past). The interior of the cone is similarly the set of all points \( q \) that have a causal curve from \( p \) to \( q \) (for the future) and \( q \) to \( p \) (for the past). The boundary is the limit of observable past and reachable future events, and one may think of the entirety of the cone as all from the point \( p \) physically accessible points.

However, all points not in the cone are in the elsewhere of \( p \), and in particular the surface \( (x,y,0) \) is the set of all points of space at the present time for \( p \). These points are of course inaccessible to \( p \), since it is impossible to travel to any other point in no time at all. We will talk more about this property of Lorentzian manifolds in chapter 5. In fact, the main reason we bring up this specific category of Lorentzian manifolds, is because the physical applications are plentiful.

3.3 Examples

In this section we look at two examples - first the 2-sphere which is a Riemannian manifold, and then a similar manifold that we choose to call the Lorentzian 2-sphere. We will see that the metric tensor, curvature tensor and christoffel symbols are all similar for the two cases; however, the geometries of these two
manifolds are very different.

**Example 3.5** (Riemannian 2-sphere). Consider a two-dimensional Riemannian manifold $\mathcal{M}$ with some coordinate system $(U, \phi = (u, v))$.

Let the metric tensor $g$ be positive definite, with components be given by:

$$g_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 u \end{bmatrix}$$

and therefore also length segment: $ds^2 = du^2 + \sin^2 u dv^2$, and metric signature $(2, 0)$.

This leads to the nonzero Christoffel symbols:

$$\Gamma^u_{vv} = -\sin u \cos u, \quad \Gamma^v_{uv} = \Gamma^v_{vu} = \cos u \sin u.$$  

Now, these coordinates are clearly orthogonal, letting us simplify the geodesic equations to the equations in Example 3.2.

In this metric, any vector fields $X$ is positive definite, that is

$$\sum_{\alpha} \sum_{\beta} g_{\alpha\beta} X^\alpha X^\beta > 0,$$

for $X$ not a zero vector, because:

$$\sum_{\alpha} \sum_{\beta} g_{\alpha\beta} X^\alpha X^\beta = \sum_{\alpha} g_{\alpha\alpha} X^\alpha = (X^u)^2 + \sin^2 u (X^v)^2.$$  

This is clearly always positive-definite, which is defining for the Riemannian geometry.

Noting that only $\partial_u g_{vv} \neq 0$, the sum in the equation disappears and we get two equations:

$$\frac{d}{ds} \left( \frac{du}{ds} \right) - \sin u \cos u \left( \frac{dv}{ds} \right)^2 = 0,$$

$$\frac{d}{ds} \left( \sin^2 u \frac{dv}{ds} \right) = 0.$$

Without attempting to solve them, we can see that a curve $\gamma(s) = (u(s), v(s)) = (s, v_0)$, $v_0$ constant, satisfies these equations and is therefore a geodesic on this Riemannian manifold.

Looking at the curvature tensor, after some tedious but simple calculations we can derive that it has, relative to $(u, v)$, four nonzero components:

$$R^u_{uvu} = -\sin^2 u, \quad R^u_{vuv} = \sin^2 u,$$

$$R^v_{uuv} = -1, \quad R^v_{uvu} = 1.$$  

We may notice that the Riemannian manifold is not flat, but has ‘a little’ curvature. This should come as no surprise, since the components of the metric tensor are not constant.

Now that we have explored many of the properties of this manifold, we will continue to the less intuitive case of a Lorentzian manifold.  

□
Example 3.6 (Lorentzian 2-sphere). Consider a two-dimensional Lorentzian manifold \( \mathcal{M} \) with some coordinate system \( \phi = (u, v) \), much like the previous Riemannian manifold.

Unlike the Riemannian manifold, however, this Lorentzian manifold has metric signature \((1, 1)\). Now, let the metric tensor \( g \) be indefinite, with components be given by:

\[
g_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & -\sin^2 u \end{bmatrix}
\]

Now, the question arises: what is the difference between these geometries? Well, it is clear that the line segment takes a different form:

\[
ds^2 = du^2 - \sin^2 u dv^2,
\]

and it is clear that some counterintuitive properties follow from allowing that minus sign! \( \square \)
4 The Einstein Field Equations

We are now ready to treat the physical applications of semi-Riemannian manifolds. Since we started discussing semi-Riemannian manifolds, we have discussed measurement of distance and curvature. In this chapter, we apply the results of these theorems and the categorizations of semi-Riemannian manifolds from chapter 3.

In this chapter, we view the physical universe as a 4-dimensional Lorentzian manifold, eventually postulating the Einstein field equations, which formulate the relation between curvature and stress-energy content in the universe. This allows us to view the curvature tensor as a physical property of the universe, as a function of mass, momentum and energy.

Solving these equations through our main theorem, we can find the metric tensor for this manifold which allows us to, for example, formulate the geodesic equations, which describe straight paths through the physical spacetime. These two theorems are fundamental for our understanding of how semi-Riemannian manifolds relate to physics, and from a mathematical perspective they can be seen as a useful and in-depth example of problems that can be formulated and solved on a semi-Riemannian manifold.

This solution that we present, the Schwarzschild solution, has many useful applications in practice, and we shall show in the next chapter that it can describe the orbits of celestial bodies with increased accuracy over the Newtonian model. However, we start by formulating the special theory of relativity in terms of semi-Riemannian geometry before generalizing.

4.1 Special relativity

The special theory of relativity is a generalization of the Newtonian mechanics and electrodynamics to high velocities. It is based upon two principles: that in all inertial reference frames the laws of physics have the same form and that in all inertial reference frames, the speed of light in a vacuum are the same.

The special theory of relativity is modelled as the four-dimensional Minkowski spacetime, \( \mathbb{M} = \mathbb{R}^1 \times \mathbb{R}^3 \). One significant aspect when studying the motion of a particle in the special theory of relativity is that and even though we still have the timelike coordinate \( t \), we cannot talk about \( t \) as some form of objective time for the particle’s movement (see e.g. [11] or [4]).

Instead, we call \( t \) coordinate time, and use a different function, \( \tau \), to denote what we call proper time. This proper time is the time as measured by the particle itself, that is, in an inertial reference frame where the particle appears to be at rest.

Now, since \( \mathbb{M} = \mathbb{R}^4 \), the metric tensor of \( \mathbb{M} \) is usually called \( \eta \), with components given by:

\[
\eta_{\alpha\beta} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix}
\]

This gives us the line element in the coordinate system \( \phi = (ct, x, y, z) \):

\[
ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2.
\]
It is from this that we define proper time. We could call it the distance from one event to another in spacetime.

**Definition 4.1** (Proper time). *Proper time* $\tau$ for a particle is the coordinate interval between points in the history of the particle. It is defined in terms of the line element (from Definition 2.13)

$$q = \sum_{\alpha} \eta_{\alpha\alpha} dx^\alpha dx^\alpha.$$ 

Since $q$ is a quadratic form, we define the proper time squared as $q$,

$$c^2 d\tau^2 = q = c^2 dt^2 - dx^2 - dy^2 - dz^2.$$

Physically speaking, this equation rewritten gives us an idea of the relationship between coordinate time and proper time for the particle:

$$d\tau = \sqrt{c^2 dt^2 - \left(\frac{dx}{dt}\right)^2 dt^2 - \left(\frac{dy}{dt}\right)^2 dt^2 - \left(\frac{dz}{dt}\right)^2 dt^2} = \sqrt{1 - \frac{v^2}{c^2}} dt,$$

where $v$ is the 3-space velocity of the particle.

Since coordinate time and proper time are not the same, it is obvious that a particle travelling with speed $v'$ relative to another particle will have a different proper time.

More significant is that the addition of velocities $v_1$ and $v_2$ is not as in Newtonian mechanics $v = v_1 + v_2$.

Neglecting to derive the equation we arrive at the result

$$v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}},$$

and it is this result that leads us to conclude that it is impossible for material particles to ever reach a speed $v \geq c$, since if $v_1 \leq c, v_2 \leq c, v \leq c$.

However, we note that if either speed is already $c$, then the resulting speed is always $c$, in correspondence with the second principle of special relativity.

### 4.1.1 Generalization of Newtonian concepts

We have spent only a little time on explaining the geometry of special relativity, and now we move on to introduce the generalized concepts that we need to understand how this special theory is supposed to replace Newtonian mechanics. First of all, we redefine velocity.

In Newtonian mechanics, we use the 3-space velocity, or coordinate velocity, $v^\alpha = dx^\alpha/dt$. However, we introduce the world velocity, or 4-velocity: $U^\alpha = du^\alpha/d\tau$, measured along the path of the particle itself.

From this we derive that the 4-force on the particle of mass $m$ is defined as:

$$F^\alpha = \frac{dU^\alpha}{d\tau} = m \frac{dU^\alpha}{d\tau}.$$

Then we can solve physical problems using this version of Newton’s second law, and under any linear coordinate transformation (Lorentz transformation) of the
coordinates $\phi = u^\alpha$, it is required from the first principle of special relativity that these laws are invariant. These transformations can be written in terms of the coordinates as

$$\tilde{u}^\alpha = \sum_\beta \Lambda^\alpha_\beta u^\beta + c^\alpha, \quad (4.1)$$

where $\Lambda$ is required to satisfy

$$\sum_\alpha \sum_\beta \eta_{\alpha\beta} \Lambda^\alpha_\mu \Lambda^\beta_\nu = \eta_{\mu\nu}.$$  

It can be shown that these laws hold up under the transformations (4.1), however we focus on the shortcomings of the special theory; namely, that the force felt from gravity by a particle is not in general the same for two different observers (see e.g. chapter 2 in [11]).

This inconsistency required the formulation of a theory that incorporates gravity into the relativistic framework.

### 4.2 General relativity

The general theory of relativity in turn generalizes the special theory of relativity, and in particular it describes the force of gravity as well as all the effects of special relativity, by using a semi-Riemannian manifold of metric signature $(1, 3)$. It models the effects of gravity as the curvature of this manifold.

It of course also generalizes the special theory of relativity by using an in general non-flat metric tensor, and in fact is required to approximate to special relativity locally. Special relativity is a special case of general relativity, where there is no gravitational force acting on the particle.

The general theory of relativity models the universe as a spacetime, a $(1, 3)$ Lorentzian manifold with a metric tensor which is not flat in general, but varies according to a few fundamental principles. The general theory is foremost a generalization of the special theory, and depends on the postulates of the special theory.

**Equivalence principle**

The equivalence principle states that for a from-infinity nonrotating small neighborhood of spacetime on which the force of gravity is the only force acting upon it, the laws of physics are the laws of special relativity.

Since there is now no distinction between a particle at rest in a gravitational field and a particle accelerated by a force equal to gravity, we can categorize the motion of objects in free fall under the influence of gravity as travelling along geodesics on the spacetime manifold.

**Energy-momentum curves spacetime**

The curvature of the Lorentzian manifold of spacetime is caused by energy-momentum. Since geodesics on this manifold are motions of particles in free fall; that is, only affected by the force of gravity, curvature and gravitation are linked.

The source of gravity is energy-momentum, and the source of curvature in this manifold is gravity. The precise equation for this relation is described later in this section, and is the foundation of general relativity.
4.2.1 The Einstein field equations

First, we have decided that energy-momentum curves spacetime. We will work with a manifold \( \mathcal{M} \) with metric signature \((1, 3)\) and metric tensor \( g \). It is obvious that the equations governing the relation between curvature and energy-momentum should contain some form of the curvature tensor and some form of tensor energy-momentum tensor (see e.g. [11]).

**Definition 4.2 (Stress-energy tensor).** The stress-energy tensor \( T \) is a symmetric \((0, 2)\) tensor field with \( \text{div} \, T = 0 \).

We choose to model the distribution of mass and momentum as a fluid, and let the components be given by:

\[
T_{\alpha\beta} = \left( \rho + \frac{p}{c^2} \right) U_{\alpha} U_{\beta} - pg_{\alpha\beta}, \tag{4.2}
\]

where \( \rho \) is density and \( p \) is pressure. \( U \) is the 4-velocity discussed in the section on special relativity.

The stress-energy tensor has vanishing divergence due to the conservation of energy and momentum, and its components describe the mass-energy and momentum content in the spacetime model.

We will now come to the most important part of the general theory: the Einstein field equations. These equations are solved for the metric tensor. Therefore, knowing the Einstein field equations, and the energy contained in the spacetime, we can know the metric tensor, and therefore the geometry of spacetime.

It should perhaps be noted that in physical applications, we usually choose a body around which we describe the curvature of the body, and approximate gravitational effects from other bodies as negligible, and therefore often get something like a spherically-symmetric manifold to work with.

**Definition 4.3 (The Einstein field equations).** Let \( \mathcal{M} \) be the semi-Riemannian manifold of metric signature \((1, 3)\) representing spacetime, with stress-energy tensor \( T \) and curvature tensor \( R \). Then, we consider the following equations:

\[
C^1_3(R) - \frac{1}{2} g C^1_1(\Gamma^1_1 C^1_3(R)) = kT, \tag{4.3}
\]

which we call the *Einstein equations*, where \( k \) is a constant and \( g \) is unknown. These equations are commonly written in component-free form using the Ricci tensor and Ricci scalar as:

\[
\text{Ric} - \frac{1}{2} g S = kT. \tag{4.3}
\]

Since the stress-energy tensor is a symmetric \((0, 2)\) tensor with vanishing divergence, we are interested in showing that the left-hand side (usually called the *Einstein tensor*) is also a symmetric \((0, 2)\) tensor with vanishing divergence. Otherwise, of course, these equalities cannot hold, and in fact Einstein first proposed a left-hand side of this equation that turned out not to have a vanishing divergence (see e.g. chapter 3 in [4]).

**Theorem 4.4 (The Einstein tensor).** *The Einstein tensor, \( G = \text{Ric} - \frac{1}{2} g S \), is a symmetric, \((0, 2)\) tensor field with vanishing divergence.*
Proof. First we prove that $G$ is a $(0,2)$ tensor field: That it is a $(0,2)$ tensor field is clear from the fact that $g$ is a $(0,2)$ tensor and from the definition of contraction, the contraction of the $(1,3)$ tensor $R$ is $C^1_2(R) \in T^2_0(M)$.

Now we prove that $G$ is symmetric: Since $g$ is symmetric and by Lemma 2.30, the Ricci tensor is symmetric, $G$ is clearly symmetric.

To prove that $\text{div} \, G = 0$: Inspired by the proof of chapter 3 in [4], we prove this by the second Bianchi identity.

The covariant derivative of the curvature tensor satisfies the second Bianchi identity, from equation (2.15):

$$\nabla_\nu R^\mu_\alpha\beta + \nabla_\beta R^\mu_\alpha\nu + \nabla_\lambda R^\mu_\alpha\beta = 0. \tag{4.4}$$

Contracting the left-hand side of this equation, we can define a tensor (equal to zero by equation (4.4)):

$$\nabla_\nu \tilde{R}_\alpha\beta = \sum_\mu \left( \nabla_\nu R^\mu_\alpha\beta + \nabla_\beta R^\mu_\alpha\nu + \nabla_\lambda R^\mu_\alpha\beta \right).$$

By the definition of the Ricci tensor, we can rewrite this as

$$\nabla_\nu \tilde{R}_\alpha\beta = \nabla_\nu \text{Ric}_\alpha\beta + \sum_\mu \left( \nabla_\beta R^\mu_{\alpha\nu} - \nabla_\nu R^\mu_{\alpha\beta} \right),$$

and by the skew-symmetry property of the curvature tensor and the definition of the Ricci tensor, we can get

$$\nabla_\nu \tilde{R}_\alpha\beta = \nabla_\nu \text{Ric}_\alpha\beta - \nabla_\beta \text{Ric}_\alpha\nu + \sum_\mu \nabla_\mu R^\mu_{\alpha\beta}.$$ 

Now, we define a new tensor by raising an index and contracting:

$$\mathcal{R} = C^1_2 \tilde{R},$$

$$\mathcal{R} = \sum_\alpha \nabla_\alpha \text{Ric}_\beta^\alpha - \nabla_\beta S + \sum_\mu \sum_\alpha \nabla_\mu R^\mu_{\alpha\beta}.$$ 

Now, we rewrite all terms on the right-hand side of this equation: by the definition of divergence, the first term is

$$\sum_\alpha \nabla_\alpha \text{Ric}_\beta^\alpha = \text{div}_\alpha \text{Ric}_\beta^\alpha.$$ 

By the definition of divergence, the second term is

$$\nabla_\beta S = \text{div}_\beta S.$$ 

Finally, the third term can be rewritten by the definition of the Ricci tensor (but note the raised index):

$$\sum_\mu \sum_\alpha \nabla_\mu R^\mu_{\alpha\beta} = \sum_\mu \nabla_\mu \text{Ric}_\beta^\mu.$$ 

38
and by the definition of the divergence, this is $\text{div}_\mu \text{Ric}_\mu$. Now, just rewriting $R$ by changing the label in the third term from $\mu$ to $\alpha$ it becomes

$$R = \sum_\alpha \nabla_\alpha \left( \text{Ric}_\beta^\alpha - \frac{1}{2} S_\delta^\alpha \right),$$

where $\delta^\alpha$ is the Kronecker delta. By the definition of divergence, this is:

$$R = \text{div}_\alpha \left( \text{Ric}_\beta^\alpha - \frac{1}{2} S_\delta^\alpha \right).$$

Now, by the definition of the Einstein tensor:

$$\text{Ric}_\beta^\alpha - \frac{1}{2} S_\delta^\alpha = \tau_1^1 \text{G}.$$

Finally, from the second Bianchi identity (4.4), we get that

$$\text{div} \tau_1^1 \text{G} = 0$$

This proves that the Einstein tensor $\text{G}$ has a vanishing divergence.

Now that we have established this relation between curvature and energy, we can delve into the solution for these equations in Theorem 4.8 in the coming section.

### 4.3 The Schwarzschild geometry

The first solution to the Einstein field equations presented was the Schwarzschild solution. It is the most general solution for the physical spacetime surrounding a static, spherically symmetric, uncharged, at infinity non-rotating massive body, such as a star with negligible rotational energy\(^1\) (see e.g. chapter 23 in [11]).

We therefore look at the 4-dimensional Lorentzian manifold $\mathcal{M}$ with the topology of $\mathbb{R}^4$. Now, we more rigorously investigate the geometric and topological properties of the Schwarzschild spacetime. It is common to define the geometry of a semi-Riemannian manifold in terms of the line element $q$ (see Definition 2.13).

1. **The physical spacetime is given to be static.**

What we mean by static is that the geometry of $\mathcal{M}$ does not depend on any timelike coordinate function.

For this reason, we will model our four-dimensional spacetime as a semi-Riemannian product manifold $\mathbb{R} \times \mathbb{R}^3$ with a metric tensor that does not depend on the timelike coordinate. Using Euclidean space (Example 2.8) as a model for our spacelike coordinates, we can write the line element $q$ as

$$q = \tilde{A}(x, y, z)c^2 dt^2 - \tilde{B}(x, y, z)dx^2 - \tilde{C}(x, y, z)dy^2 - \tilde{D}(x, y, z)dz^2,$$

\(^1\)The spacetime outside of our own star is a good approximation to this kind of spacetime, even though the sun has some rotational energy.
where $c$ is a constant, $\tilde{A}: \mathbb{R}^3 \to \mathbb{R}$, $\tilde{B}: \mathbb{R}^3 \to \mathbb{R}$, $\tilde{C}: \mathbb{R}^3 \to \mathbb{R}$, $\tilde{D}: \mathbb{R}^3 \to \mathbb{R}$.

For a static spacetime, we can separate timelike and spacelike coordinates. Therefore we see the topology of this spacetime as a product manifold of $\{t: t \in \mathbb{R}\}$ and the spacelike 3-plane $(x, y, z), \mathbb{R} \times \mathbb{R}^3$.

2. The physical spacetime is spherically symmetric.

The spacetime is spherically symmetric. For the topology, this means that the topology of the spacelike coordinates is isometric to the 2-sphere as $\mathbb{R}^3$ is diffeomorphic to $\mathbb{R}^+ \times S^2$. For the geometry, this means that the geometry does not depend on the coordinates of the 2-sphere. In terms of the line element $q$, we can write this

$$q = A(R)c^2 dt^2 + B(R)dR^2 + R^2 d\Omega^2,$$

where $A: \mathbb{R}^+ \to \mathbb{R}$, $B: \mathbb{R}^+ \to \mathbb{R}$ and $R^2 d\Omega = R^2 d\theta^2 + R^2 \sin^2 \theta d\phi^2$, in terms of the 2-sphere coordinates (Example 2.10).

This symmetry gives that for each $R$, $\mathbb{R}^+ \times S^2$ has the geometry of a 2-sphere of radius $R$. $R^2 d\Omega^2$ can be thought of as the surface line element of this product, for each $R$ (to understand this spherical symmetry in more detail, see Example 7.37 and chapter 13 in [16]).

3. The physical spacetime is flat at infinity.

What we mean by flat at infinity is that, $\lim_{R \to \infty} \text{Ric}_{\alpha\beta} = 0$ for all $\alpha, \beta$, and the line element reduces to, in the topology of $(U, \phi = (ct, R, \theta, \varphi))$

$$q = c^2 dt^2 - dR^2 - R^2 d\Omega^2.$$

From these properties, we have come to a unique solution under the conditions imposed in the beginning of this section, but first we will need some supporting lemmas (see e.g. chapter 13 in [16] for a similar perspective on these properties).

**Lemma 4.5** (Ricci tensor components in the Schwarzschild solution). For a 4-dimensional spherically symmetric Lorentzian manifold $\mathcal{M}$ with metric tensor

$$g_{\alpha\beta} = \begin{bmatrix} e^{2f_1(R)} & 0 & 0 & 0 \\ 0 & -e^{2f_2(R)} & 0 & 0 \\ 0 & 0 & -R^2 & 0 \\ 0 & 0 & 0 & -R^2 \sin^2 \theta \end{bmatrix}$$

relative to a coordinate system $(U, \phi = (ct, R, \theta, \varphi))$, where $c$ is a constant, the Ricci tensor components are given by

$$\text{Ric}_{tt} = \frac{d^2 f_1}{dR^2} + \left(\frac{df_1}{dR}\right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_1}{dR} \left(e^{2f_1-f_2}\right),$$

$$\text{Ric}_{RR} = -\frac{d^2 f_1}{dR^2} + \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_2}{dR} - \left(\frac{df_1}{dR}\right)^2,$$

$$\text{Ric}_{\theta\theta} = -e^{-2f_2} + R e^{-2f_2} \frac{df_2}{dR} - \frac{df_1}{dR} R e^{-2f_2} + 1,$$

$$\text{Ric}_{\varphi\varphi} = \text{Ric}_{\theta\theta} \sin^2 \theta.$$
Proof. Using this metric tensor, we start by noticing that it is orthogonal. This means that the Christoffel symbol components (2.8) can be written as

\[ \Gamma^\lambda_{\alpha\beta} = \frac{1}{2} g^{\lambda\kappa} (\partial_\alpha g_{\kappa\beta} + \partial_\beta g_{\kappa\alpha} - \partial_\kappa g_{\alpha\beta}). \]

This gives us nine independent nonzero Christoffel symbols,

\[ \Gamma^R_{\theta\theta} = -R \sin^2 \theta \rho e^{-2f_z}, \quad \Gamma^R_{\rho\theta} = \Gamma^R_{\theta\rho} = R^{-1}, \quad \Gamma^\theta_{\varphi\varphi} = -\sin \theta \cos \theta, \quad \Gamma^\varphi_{R\varphi} = \Gamma^\varphi_R = R^{-1}, \quad \Gamma^\varphi_{\varphi\theta} = \cos \theta \sin \theta. \]

From Theorem 2.27 and the definition of the Ricci tensor, the components of the Ricci tensor are given by

\[ \text{Ric}_{\alpha\beta} = \sum_\mu R^\mu_{\alpha\beta\mu} = \sum_\mu \left( \partial_\mu \Gamma^\mu_{\beta\alpha} - \partial_\beta \Gamma^\mu_{\mu\alpha} + \sum_\nu \Gamma^\nu_{\mu\nu} \Gamma^\mu_{\beta\nu} - \sum_\nu \Gamma^\mu_{\beta\nu} \Gamma^\nu_{\mu\alpha} \right). \]

This gives us four nonzero Ricci tensor components,

\[ \text{Ric}_{tt} = \sum_\mu \left( \partial_t \Gamma^\mu_{tt} - \partial_t \Gamma^\mu_{t\mu} + \sum_\nu \Gamma^\nu_{\mu\nu} \Gamma^\mu_{tt} - \sum_\nu \Gamma^\nu_{tt} \Gamma^\mu_{\mu\nu} \right), \quad (4.10) \]

\[ \text{Ric}_{RR} = \sum_\mu \left( \partial_\mu \Gamma^\mu_{RR} - \partial_R \Gamma^\mu_{\mu R} + \sum_\nu \Gamma^\nu_{\mu\nu} \Gamma^\mu_{RR} - \sum_\nu \Gamma^\nu_{RR} \Gamma^\mu_{\mu\nu} \right), \quad (4.11) \]

\[ \text{Ric}_{\theta\theta} = \sum_\mu \left( \partial_\mu \Gamma^\mu_{\theta\theta} - \partial_\theta \Gamma^\mu_{\mu \theta} + \sum_\nu \Gamma^\nu_{\mu\nu} \Gamma^\mu_{\theta\theta} - \sum_\nu \Gamma^\nu_{\theta\theta} \Gamma^\mu_{\mu\nu} \right), \quad (4.12) \]

\[ \text{Ric}_{\varphi\varphi} = \sum_\mu \left( \partial_\mu \Gamma^\mu_{\varphi\varphi} - \partial_\varphi \Gamma^\mu_{\mu \varphi} + \sum_\nu \Gamma^\nu_{\mu\nu} \Gamma^\mu_{\varphi\varphi} - \sum_\nu \Gamma^\nu_{\varphi\varphi} \Gamma^\mu_{\mu\nu} \right). \quad (4.13) \]

First we derive equation (4.10). It becomes in terms of \( \Gamma^\alpha_{\alpha\beta} \) (only writing out the nonzero terms)

\[ \text{Ric}_{tt} = \partial_R \Gamma^R_{tt} + \sum_\mu \Gamma^\mu_{t\mu} \Gamma^R_{tt} - \sum_\mu \sum_\nu \Gamma^\nu_{tt} \Gamma^\mu_{\mu\nu}. \quad (4.14) \]

Notice that the last term is only nonzero when \( \mu = t, \nu = R \) and vice versa. The second term in (4.14) is nonzero for all \( \mu \), so we get

\[ \text{Ric}_{tt} = \partial_R \Gamma^R_{tt} + \Gamma^t_{tt} \Gamma^R_{tt} + \Gamma^R_{tt} \Gamma^R_{tt} + \Gamma^R_{tt} \Gamma^R_{tt} + \Gamma^R_{tt} \Gamma^R_{tt} - 2 \Gamma^t_{tt} \Gamma^R_{tt}, \]

or equivalently,

\[ \text{Ric}_{tt} = \partial_R \left( \frac{df_1}{dR} e^{2(f_z - f_z)} \right) + \left( \frac{df_1}{dR} + \frac{df_2}{dR} + R^{-1} + R^{-1} - 2 \frac{df_1}{dR} \right) \frac{df_1}{dR} e^{2(f_z - f_z)}, \]

and therefore

\[ \text{Ric}_{tt} = \frac{df_1}{dR} e^{2(f_z - f_z)} + 2 \frac{df_1}{dR} e^{2(f_z - f_z)} \left( \frac{df_1}{dR} - \frac{df_2}{dR} \right) \left( \frac{df_1}{dR} - \frac{df_2}{dR} + 2 \frac{df_1}{dR} \right) e^{2(f_z - f_z)}, \]

41
or simply
\[
\text{Ric}_{tt} = \left( \frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_1}{dR} \right) e^{2(f_1-f_2)}.
\]

This proves that the $tt$-component of the Ricci tensor is equal to (4.6).

Now we derive equation (4.11). In terms of $\Gamma^\lambda_{\alpha\beta}$ (again only writing out the nonzero terms) we have
\[
\text{Ric}_{RR} = \partial_R \Gamma^R_{RR} - \partial_R \left( \Gamma^t_{tR} + \Gamma^R_{RR} + \Gamma^\varphi_{R\varphi} \right) \right)
+ \sum_\mu \sum_\nu \Gamma^\mu_{\mu\nu} \Gamma^\nu_{RR} - \sum_\mu \Gamma^\mu_{R\nu \Gamma^\nu_{RR}}.
\]

Clearly the third term is only nonzero for $\nu = R$. The fourth term is only nonzero for $\mu = \nu$. This gives us
\[
\text{Ric}_{RR} = \partial_R \Gamma^R_{RR} - \partial_R \left( \Gamma^t_{tR} + \Gamma^R_{RR} + \Gamma^\varphi_{R\varphi} \right) \right)
+ \Gamma^R_{tR} \Gamma^R_t + \Gamma^R_{RR} \Gamma^R_R + \Gamma^\theta_{RR} \Gamma^R_{RR} + \Gamma^\varphi_{R\varphi} \Gamma^\varphi_{R\varphi}.
\]

or in terms of the metric tensor components,
\[
\text{Ric}_{RR} = \frac{d^2 f_2}{dR^2} - \partial_R \left( \frac{df_1}{dR} + \frac{df_2}{dR} + R^{-1} + R^{-1} \right)
\]
\[
+ \left( \frac{df_1}{dR} + \frac{df_2}{dR} + R^{-1} + R^{-1} \right) \frac{df_2}{dR}
\]
\[
- \left( \left( \frac{df_1}{dR} \right)^2 + \left( \frac{df_2}{dR} \right)^2 + R^{-2} + R^{-2} \right).
\]

or equivalently,
\[
\text{Ric}_{RR} = - \frac{d^2 f_1}{dR^2} + \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_2}{dR} - \left( \frac{df_1}{dR} \right)^2.
\]

This concludes the proof of equation (4.7).

Next, we shall derive equation (4.12). Again in terms of the non-zero Christoffel symbols, we get:
\[
\text{Ric}_{\theta\theta} = \partial_R \Gamma^\theta_{\theta\theta} - \partial_\theta \Gamma^\varphi_{\varphi\theta} + \sum_\mu \left( \sum_\nu \Gamma^\mu_{\mu\nu} \Gamma^\nu_{\theta\theta} - \sum_\nu \Gamma^\mu_{\theta\theta} \Gamma^\nu_{\mu\nu} \right).
\]

In the third term, clearly only for $\nu = R$ is the term nonzero. The fourth term is nonzero in three cases: for $\mu = \theta$, $\nu = R$ (and vice versa), and for $\mu = \nu = \varphi$. Considering this, the previous equation becomes
\[
\text{Ric}_{\theta\theta} = \partial_R \Gamma^\theta_{\theta\theta} - \partial_\theta \Gamma^\varphi_{\varphi\theta} + \left( \Gamma^t_{tR} + \Gamma^R_{RR} + \Gamma^\theta_{R\theta} + \Gamma^\varphi_{R\varphi} \right) \Gamma^R_{\theta\theta} - \left( \Gamma^\varphi_{\theta\theta} \right)^2 - 2\Gamma^\varphi_{R\theta} \Gamma^R_{\theta\theta}.
\]

42
In terms of the metric tensor, this can be written

\[ \text{Ric}_{\theta\theta} = -\partial_{\theta} \left( R e^{-2f_2} \right) - \partial_{\theta} \left( \frac{\cos \theta}{\sin \theta} \right) - \left( \frac{df_1}{dR} + \frac{df_2}{dR} + R^{-1} + R^{-1} \right) R e^{-2f_2} \]

\[ - \left( \frac{\cos \theta}{\sin \theta} \right)^2 + 2e^{-2f_2}, \]

or equivalently, after some manipulation,

\[ \text{Ric}_{\theta\theta} = -e^{-2f_2} + 2R e^{-2f_2} \frac{df_2}{dR} - \left( \frac{df_1}{dR} + \frac{df_2}{dR} \right) R e^{-2f_2} - 2e^{-2f_2} + 1 + 2e^{-2f_2}, \]

and

\[ \text{Ric}_{\theta\theta} = -e^{-2f_2} + R e^{-2f_2} \frac{df_2}{dR} - \frac{df_1}{dR} R e^{-2f_2} + 1. \]

This is clearly equal to (4.8).

To derive equation (4.13). In terms of the non-zero Christoffel symbols, we have

\[ \text{Ric}_{\varphi\varphi} = \partial_{\varphi} \Gamma_R^{\varphi} + \partial_{\varphi} \Gamma_{\varphi\varphi} + \sum_{\mu} \left( \sum_{\nu} \Gamma_{\mu\nu}^{\varphi} \Gamma_{\varphi\varphi}^{\nu} - \sum_{\nu} \Gamma_{\varphi\nu}^{\varphi} \Gamma_{\nu\varphi}^{\varphi} \right) \].

Now we look at the separate terms. The third term is nonzero in two cases: for \( \nu = R \) and for \( \nu = \theta \). The fourth term is nonzero in two cases: for \( \nu = \varphi \), \( \mu = \theta \) (and vice versa), as well as for \( \nu = \varphi \), \( \mu = R \) (and vice versa). Then this equation becomes

\[ \text{Ric}_{\varphi\varphi} = \partial_{\varphi} \Gamma_R^{\varphi} + \partial_{\varphi} \Gamma_{\varphi\varphi} + \left( \Gamma_{\varphi}^R + \Gamma_{R\varphi}^R + \Gamma_{\theta\varphi}^R + \Gamma_{\varphi\varphi}^\varphi \right) \Gamma_{\varphi\varphi}^{\varphi} + \Gamma_{\varphi\varphi}^\varphi \Gamma_{\varphi\varphi}^{\varphi} \]

\[ - 2 \Gamma_{\varphi\varphi}^{\varphi} \Gamma_{\theta\varphi}^\varphi - 2 \Gamma_{\varphi\varphi}^R \Gamma_{R\varphi}^{\varphi}, \]

or written in terms of the metric tensor,

\[ \text{Ric}_{\varphi\varphi} = \partial_{\varphi} \left( -R \sin^2 \theta e^{-2f_2} \right) - \partial_{\theta} (\sin \theta \cos \theta) \]

\[ - \left( \frac{df_1}{dR} + \frac{df_2}{dR} + R^{-1} + R^{-1} \right) R \sin^2 \theta e^{-2f_2} - \frac{\cos \theta}{\sin \theta} \sin \theta \cos \theta \]

\[ + 2 \sin \theta \cos \theta \frac{\cos \theta}{\sin \theta} + 2R \sin^2 \theta e^{-2f_2} R^{-1}, \]

which can be rewritten

\[ \text{Ric}_{\varphi\varphi} = -\sin^2 \theta e^{-2f_2} + 2R \sin^2 \theta e^{-2f_2} \frac{df_2}{dR} - \left( \frac{df_1}{dR} + \frac{df_2}{dR} \right) R \sin^2 \theta e^{-2f_2} + \sin^2 \theta, \]

or,

\[ \text{Ric}_{\varphi\varphi} = -\sin^2 \theta e^{-2f_2} + R \sin^2 \theta e^{-2f_2} \frac{df_2}{dR} - R \sin^2 \theta e^{-2f_2} \frac{df_1}{dR} + \sin^2 \theta, \]

which is actually \( \text{Ric}_{\theta\theta} \sin^2 \theta \).
Lemma 4.6 (Scalar curvature of the Schwarzschild manifold). For a 4-dimensional spherically symmetric Lorentzian manifold $\mathcal{M}$ with metric tensor

$$g_{\alpha\beta} = \begin{bmatrix} e^{2f_1(R)} & 0 & 0 & 0 \\ 0 & -e^{2f_2(R)} & 0 & 0 \\ 0 & 0 & -R^2 & 0 \\ 0 & 0 & 0 & -R^2 \sin^2\theta \end{bmatrix}$$

(4.15)

relative to a coordinate system $(\mathcal{U}, \phi = (ct, R, \theta, \varphi))$, where $c$ is a constant, the scalar curvature is given by

$$S = 2e^{-2f_2} \left( \frac{d^2f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + R^{-2} - R^{-2}e^{2f_2} \right).$$

Proof. By Definition 2.31, $S$ is the contraction of the raised-index Ricci tensor, it can be written in terms of the Ricci components as

$$S = \sum_\alpha \sum_\beta g^{\alpha\beta} \text{Ric}_{\alpha\beta} = \sum_\alpha g^{\alpha\alpha} \text{Ric}_{\alpha\alpha}.$$

This becomes, using the Ricci components from Lemma 4.5,

$$S = \left( \left( \frac{d^2f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} \right) e^{2(f_1 - f_2)} \right) e^{-2f_1}$$

$$- \left( - \frac{d^2f_1}{dR^2} + \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1)}{dR} \right) e^{-2f_2}$$

$$- \left( -e^{-2f_2} + Re^{-2f_2} \frac{df_2}{dR} \right) e^{-2f_1} Re^{-2f_2} + 1 \right) R^{-2}$$

$$- \left( -e^{-2f_2} + Re^{-2f_2} \frac{df_2}{dR} - \frac{df_1}{dR} \right) \sin^2\theta R^{-2} \frac{1}{\sin^2\theta},$$

or equivalently

$$S = \left( \frac{d^2f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1)}{dR} + \frac{d^2f_1}{dR^2} - \frac{df_1}{dR} \frac{df_2}{dR} - 2R^{-1} \frac{df_2}{dR} \right)$$

$$+ \left( \frac{df_1}{dR} \right)^2 e^{-2f_2} - 2 \left( -e^{-2f_2} + Re^{-2f_2} \frac{df_2}{dR} - \frac{df_1}{dR} \right) Re^{-2f_2} + 1 \right) R^{-2}.$$

Hence, removing the parentheses,

$$S = 2e^{-2f_2} \left( \frac{d^2f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + 2R^{-1} \frac{df_1}{dR} e^{-2f_2} \right.$$

$$- 2R^2 \frac{df_2}{dR} e^{-2f_2} + 2e^{-2f_2} R^{-2} - 2e^{-2f_2} \frac{df_2}{dR} R^{-1} + 2 \frac{df_1}{dR} e^{-2f_2} R^{-1} - R^{-2},$$

or, taking out a factor $2e^{-2f_2}$,

$$S = 2e^{-2f_2} \left( \frac{d^2f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + R^{-2} - R^{-2} e^{2f_2} \right).$$

This is clearly the scalar curvature that we were looking for. \hfill \Box
Lemma 4.7 (Einstein tensor components in the Schwarzschild solution). For a 4-dimensional spherically symmetric Lorentzian manifold \( M \) with metric tensor
\[
 g_{\alpha\beta} = \begin{bmatrix}
 e^{2f_1(R)} & 0 & 0 & 0 \\
 0 & -e^{2f_2(R)} & 0 & 0 \\
 0 & 0 & -R^2 & 0 \\
 0 & 0 & 0 & -R^2 \sin^2 \theta
\end{bmatrix}
\] (4.16)
relative to a coordinate system \((U, \phi = (ct, R, \theta, \varphi))\), where \( c \) is a constant, the Einstein tensor components are given by
\[
 G_{tt} = (2R \frac{df_1}{dR} + 1 + e^{2f_2}) R^{-2} e^{2(f_1 - f_2)}, \\
 G_{RR} = (2R \frac{df_1}{dR} - 1 - e^{2f_2}) R^{-2}, \\
 G_{\theta\theta} = \left( \frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + R^{-1} \frac{d(f_1 - f_2)}{dR} \right) R^2 e^{-2f_2}, \\
 G_{\varphi\varphi} = G_{\theta\theta} \sin^2 \theta.
\]

Proof. By Definition 4.3,
\[
 G_{\alpha\beta} = \text{Ric}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} S.
\]
Then, using Lemma 4.5, we get four nonzero components of the Einstein tensor,
\[
 G_{tt} = \text{Ric}_{tt} - \frac{1}{2} g_{tt} S \tag{4.17} \\
 G_{RR} = \text{Ric}_{RR} - \frac{1}{2} g_{RR} S \tag{4.18} \\
 G_{\theta\theta} = \text{Ric}_{\theta\theta} - \frac{1}{2} g_{\theta\theta} S \tag{4.19} \\
 G_{\varphi\varphi} = \text{Ric}_{\varphi\varphi} - \frac{1}{2} g_{\varphi\varphi} S. \tag{4.20}
\]

Now, expanding \( G_{tt} \) by Lemmas 4.5 and 4.6, we find
\[
 G_{tt} = \left( \frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_1}{dR} \right) e^{2(f_1 - f_2)} \\
 - \frac{1}{2} e^{2f_1} e^{-2f_2} \left( \frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} \right) \\
 + 2R^{-1} \frac{d(f_1 - f_2)}{dR} - R^{-2} - R^{-2} e^{2f_2},
\]
or equivalently,
\[
 G_{tt} = \left( \frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_1}{dR} - \frac{d^2 f_1}{dR^2} - \left( \frac{df_1}{dR} \right)^2 \\
 + \frac{df_1}{dR} \frac{df_2}{dR} - 2R^{-1} \frac{d(f_1 - f_2)}{dR} - R^{-2} + R^{-2} e^{2f_2} \right) e^{2(f_1 - f_2)},
\]
which after some simple cancellation becomes

\[ G_{tt} = \left( 2R^{-1} \frac{df_1}{dR} - 2R^{-1} \frac{d(f_1 - f_2)}{dR} - R^{-2} + R^{-2} e^{2f_2} \right) e^{2(f_1 - f_2)}, \]

which can be rewritten

\[ G_{tt} = \left( 2R \frac{df_2}{dR} - 1 + e^{2f_2} \right) R^{-2} e^{2(f_1 - f_2)}. \]

Which is the \( tt \)-component of the Einstein tensor that we set out to derive.

Now we expand \( G_{RR} \) in the same way, using the same lemmas, and we get

\[ G_{RR} = \left( -\frac{d^2 f_1}{dR^2} + \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{df_2}{dR} - \left( \frac{df_1}{dR} \right)^2 \right) \]

\[ + \frac{1}{2} e^{2f_2} e^{2f_2} \left( \frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} \right) \]

\[ + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + R^{-2} - R^{-2} e^{2f_2} \),

again simplifying, and noticing that the factor outside the second parentheses simplifies to unity, we can start cancelling terms, giving us

\[ G_{RR} = \left( 2R^{-1} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + R^{-2} - R^{-2} e^{2f_2} \right). \]

This equation is clearly equivalent to

\[ G_{RR} = \left( 2R \frac{df_1}{dR} + 1 - e^{2f_2} \right) R^{-2}. \]

This is the equation for the \( RR \)-component of \( G \).

In a similar manner as with the previous equations, we continue with \( G_{\theta \theta} \) and get

\[ G_{\theta \theta} = -e^{-2f_2} + R^{-2} e^{2f_2} \frac{df_2}{dR} - \frac{df_1}{dR} R e^{-2f_2} + 1 - \frac{1}{2} R^2 e^{-2f_2} \left( \frac{d^2 f_1}{dR^2} \right) \]

\[ + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + R^{-2} - R^{-2} e^{2f_2} \),

or, factoring out \( R^2 e^{-2f_2} \) we can rewrite this as

\[ G_{\theta \theta} = \left( -R^{-2} + R^{-1} \frac{df_2}{dR} - \frac{df_1}{dR} R^{-1} + R^{-2} e^{2f_2} - \frac{d^2 f_1}{dR^2} \right) \]

\[ + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + 2R^{-1} \frac{d(f_1 - f_2)}{dR} + R^{-2} - R^{-2} e^{2f_2} \right) R^2 e^{-2f_2},\]
\[
G_{\theta\theta} = \left( -\frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} \frac{df_2}{dR} + R^{-1} \frac{d(f_1 - f_2)}{dR} \right) R^2 e^{-2f_2}.
\]

Now we have arrived at the sought equation for \( G_{\theta\theta} \).

Finally, we see that
\[
G_{\phi\phi} = \text{Ric}_{\phi\phi} - \frac{1}{2} g_{\phi\phi} S = \text{Ric}_{\theta\theta} \sin^2 \theta - \frac{1}{2} g_{\theta\theta} \sin^2 \theta S = G_{\theta\theta} \sin^2 \theta.
\]

Which proves that the Einstein tensor components are satisfied.

Now, after some long technical lemmas, we are ready to delve into the main theorem of this section, the Schwarzschild solution to the Einstein field equations. This is the last theorem of this essay, and as we have previously discussed, it has a lot of physical significance as well as being mathematically interesting.

**Theorem 4.8 (Schwarzschild solution).** Let \((\mathfrak{M}, g)\) be a 4-dimensional Lorentzian manifold that is spherically symmetric (as discussed in the beginning of this section). As described earlier, the manifold has the topology of the product \(\mathbb{R} \times \mathbb{R} \times S^2\) and therefore we can use a chart \((U, \phi = (ct, R, \theta, \varphi))\), where \(c\) is a constant and \(\theta, \varphi\) are the 2-sphere coordinates from Example 2.10. Let \(R^*\) be a real-valued constant.

Assume that \((\mathfrak{M}, g)\) is static, and the stress-energy tensor takes the form:
\[
T_{\alpha\beta} = \begin{bmatrix}
\rho c^2 g_{tt} & 0 & 0 & 0 \\
0 & -p g_{RR} & 0 & 0 \\
0 & 0 & -p g_{\theta\theta} & 0 \\
0 & 0 & 0 & -p g_{\varphi\varphi}
\end{bmatrix},
\]

where \(p\) and \(\rho\) are functions of \(R\) defined by \(p = 0\) for \(R > R^*\) and
\[
M = \int_0^R 4\pi R^2 \rho(r)dr
\]

where \(M\) is a constant for \(R > R^*\). Finally, for \(R \to \infty\), the metric tensor of \(\mathfrak{M}\) takes the form:
\[
g_{\alpha\beta} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -R^2 & 0 \\
0 & 0 & 0 & -R^2 \sin^2 \theta
\end{bmatrix}
\]

The Einstein field equations in (4.3) have a unique solution given by:
\[
g_{\alpha\beta} = \begin{bmatrix}
1 - \frac{r_S}{R} & 0 & 0 & 0 \\
0 & -\left(1 - \frac{r_S}{R}\right)^{-1} & 0 & 0 \\
0 & 0 & -R^2 & 0 \\
0 & 0 & 0 & -R^2 \sin^2 \theta
\end{bmatrix} \quad (4.21)
\]

where \(r_S\) is a constant. For more details on \(r_S\), see Corollary 4.9.
Proof. Inspired by the schematic proof in chapter 23 in [11], we prove this theorem.

Using the coordinate system \( \phi = (ct, R, \theta, \varphi) \), we can deduce that due to spherical symmetry (which we have talked about earlier in this section), we need only consider the following ansatz:

\[
g_{\alpha\beta} = \begin{bmatrix}
e^{2f_1(R)} & 0 & 0 & 0 \\
0 & -e^{2f_2(R)} & 0 & 0 \\
0 & 0 & -R^2 & 0 \\
0 & 0 & 0 & -R^2 \sin^2 \theta
\end{bmatrix}
\] (4.22)

Since \( \partial_t \) needs to be timelike everywhere and \( \partial_R \) needs to be spacelike everywhere, due to \( \mathcal{M} \) being a Lorentzian manifold.

For some functions \( f_1: \mathcal{U} \to \mathbb{R} \) and \( f_2: \mathcal{U} \to \mathbb{R} \). An exponential function exists on every semi-Riemannian manifold (see, e.g. chapter 3 in [16]).

Now from the definition of the Einstein field equations (4.3) we have that \( G_{\alpha\beta} = kT_{\alpha\beta} \). This and Lemma 4.7 gives us a set of four equations:

\[
\begin{align*}
(2R \frac{df_1}{dR} - 1 + e^{2f_2}) R^{-2} e^{2(f_1 - f_2)} &= kc^2 f_1 \rho c^2, & (4.23) \\
2R \frac{df_1}{dR} + 1 - e^{2f_2} R^{-2} &= -ke^{2f_2} p, & (4.24) \\
-\frac{d^2 f_1}{dR^2} + \left( \frac{df_1}{dR} \right)^2 - \frac{df_1}{dR} R^{-1} \frac{d(f_1 - f_2)}{dR} + R^2 e^{-2f_2} &= -kR^2 p, & (4.25) \\
G_{\theta\theta} \sin^2 \theta &= -kR^2 \sin^2 \theta p. & (4.26)
\end{align*}
\]

We shall first solve the first of these first-order ordinary differential equations (4.23). First, we rewrite the equation as

\[
\left(2R \frac{df_2}{dR} e^{-2f_2} - e^{-2f_2} + 1\right) R^{-2} = kc^2
\]

and then this can be rewritten

\[
\frac{d}{dR} \left( R \left( 1 - e^{-2f_2} \right) \right) = kR^2 \rho c^2.
\]

Now, by putting

\[
r_S = R - Re^{-2f_2},
\]

we obtain:

\[
\frac{dr_S}{dR} = kR^2 \rho c^2;
\]

and by integrating this becomes

\[
r_S(R) = \frac{k\ell^2}{4\pi} \int_0^R 4\pi R^2 \rho dR,
\]

48
And solving this equation for \( R > R^* \), by the definition of \( \rho \), we have that \( r_S = Mkc^2/4\pi \) is a constant, and in terms of the function \( g_{RR} = e^{2f_2} \):

\[
e^{2f_2} = (1 - r_S/R)^{-1}. \tag{4.27}
\]

By the second first-order ordinary differential equation, equation (4.24), using again our \( r_S \), and equation (4.27) we obtain

\[
R^{-2} \left( 1 + 2R \frac{df_1}{dR} - (1 - r_S/R)^{-1} \right) = -k(1 - r_S/R)^{-1} p,
\]
or solved for \( df_1/dR \):

\[
\frac{df_1}{dR} = \frac{-kR^2 (1 - r_S/R)^{-1} p - 1 + (1 - r_S/R)^{-1}}{2R},
\]

and integrating, noting that \( p = 0 \) for \( R > R^* \), this becomes

\[
f_1 = \frac{1}{2} \int_0^R \left( \frac{1}{R - r_S} - \frac{1}{R} \right) dR = \frac{1}{2} \ln \left( 1 - \frac{r_S}{R} \right),
\]

which gives us the metric tensor component \( g_{tt} = e^{2f_1} \)

\[
e^{2f_1} = e^{\ln(1 - r_S/R)} = 1 - \frac{r_S}{R}
\]

This lets us rewrite equation (4.22) as:

\[
g_{\alpha \beta} = \begin{bmatrix}
1 - r_S/R & 0 & 0 & 0 \\
0 & -(1 - r_S/R)^{-1} & 0 & 0 \\
0 & 0 & -R^2 & 0 \\
0 & 0 & 0 & -R^2 \sin^2 \theta
\end{bmatrix}
\]

Which are the metric tensor components of the Schwarzschild solution (5.1).

**Corollary 4.9** (Schwarzschild radius). The Schwarzschild radius is given by

\[
r_S = \frac{2GM}{c^2}.
\]

**Proof.** By the proof of Theorem 4.8, \( r_S = Mkc^2/4\pi \), and by compatibility with the Newtonian limit, \( k = GM/c^4 \). This gives us \( r_S = GM/c^2 \).
Applications to the Schwarzschild solution

Now that we have familiarized ourselves with the Einstein field equations and the Schwarzschild solution, we are ready to look at a few interesting examples of applications. In particular, we shall in this chapter deal with geodesics on the Schwarzschild manifold, as well as the Schwarzschild black hole.

From a physical perspective the geodesic paths correspond to the motion of particles in free fall. From a mathematical perspective, the spherical symmetry of the Schwarzschild geometry leads to special classes of geodesics. The Schwarzschild black hole, however, is a Schwarzschild manifold with $r > R^*$. This may not appear to be much of a problem initially, but as we shall see, it leads to problems with the nondegeneracy of the metric tensor as $R$ approaches $r_S$.

5.1 The Schwarzschild manifold

The Schwarzschild manifold is the 4-dimensional Lorentzian manifold $\mathcal{M}$ that has a metric tensor

$$
\begin{pmatrix}
(1 - \frac{R_S}{R}) & 0 & 0 & 0 \\
0 & -(1 - \frac{R_S}{R})^{-1} & 0 & 0 \\
0 & 0 & -R^2 & 0 \\
0 & 0 & 0 & -R^2 \sin^2 \theta
\end{pmatrix}
$$

(5.1)

where $r_S$ is the Schwarzschild radius from Corollary 4.9.

This manifold (both the topology and the metric tensor) are used in the following sections on geodesics and singularities. This semi-Riemannian is called the Schwarzschild manifold, its metric tensor is called the Schwarzschild solution, or Schwarzschild metric, since it is the unique solution to the Einstein field equations under the conditions of Theorem 4.8.

Often the line element (5.2) is referred to as the Schwarzschild metric, even though it is only related to it.

5.2 Schwarzschild geodesics

We are interested in geodesics, because they are paths for particles in free fall about the massive body. These paths are good approximations of the paths of particles around most celestial bodies, for example the orbit of the Moon about the Earth, the path of a comet about the Sun or a satellite in orbit.

Remembering that these coordinates, for example the radius $R$, is not the radius we would measure along some radial path. Rather, rewriting the Schwarzschild solution in terms of the line element, we obtain (inspired by the similar derivation of geodesics in chapter 13 in [16]):

$$
c^2 d\tau^2 = \left(1 - \frac{r_S}{R}\right) c^2 dt^2 - \frac{1}{1 - \frac{r_S}{R}} dR^2 - R^2 d\theta^2 - R^2 \sin^2 \theta d\varphi^2,
$$

(5.2)
and this suggests that we would measure the proper length along a radial path as

\[ L = \int (1 - r_S/R)^{-1/2} dR. \]

Therefore we will derive the solutions to these curves and talk about the physical properties of them. First off, we will assume that the path is initially equatorial, that is, \( \theta(0) = \pi/2 \). Due to spherical symmetry, any geodesic can be reparametrized to be initially equatorial. We get four equations, but we start with the three simple equations for \( \varphi, \theta \) and \( t \):

We will, on the Schwarzschild manifold \( M \), use the Schwarzschild coordinates \( \phi = (ct, R, \theta, \varphi) \). These coordinates are orthogonal, and so we can use the geodesic equations derived in Example 2.23.

\[
\begin{align*}
\frac{d^2}{d\tau^2} \left( (1 - r_S/R) c^2 \frac{dt}{d\tau} \right) &= \frac{1}{2} \sum \partial_t g_{\alpha\alpha} \left( \frac{du^\alpha}{d\tau} \right)^2 = 0, \\
\frac{d^2}{d\tau^2} \left( R^2 \sin^2 \theta \frac{d\varphi}{d\tau} \right) &= \frac{1}{2} \sum \partial_\varphi g_{\alpha\alpha} \left( \frac{du^\alpha}{d\tau} \right)^2 = 0, \\
\frac{d^2}{d\tau^2} \left( R^2 \frac{d\theta}{d\tau} \right) &= \sin \theta \cos \theta \left( \frac{d\varphi}{d\tau} \right)^2,
\end{align*}
\]

and expanding the two first equations, we find constants for the energy\(^1\) and angular momentum of the particle. Calling the energy \( p_t \) and the angular momentum \( p_\varphi \), we obtain:

\[ \left( 1 - \frac{r_S}{R} \right) \frac{dt}{d\tau} = p_t, \]
\[ R^2 \frac{d\varphi}{d\tau} = p_\varphi, \]
\[ \theta = \pi/2. \]

From the line element (5.2), we get an energy equation (remembering \( \theta = \pi/2 \)):

\[ c^2 = \left( 1 - \frac{r_S}{R} \right) c^2 \left( \frac{dt}{d\tau} \right)^2 - \frac{1}{1 - \frac{r_S}{R}} \left( \frac{dR}{d\tau} \right)^2 - R^2 \left( \frac{d\varphi}{d\tau} \right)^2, \]

and therefore

\[ \left( \frac{dR}{d\tau} \right)^2 = c^2 p_t^2 - \left( 1 - \frac{r_S}{R} \right) p_\varphi^2 + c^2 r_S - c^2. \quad (5.3) \]

Now we can see the last three terms on the right-hand side as a sort of potential energy of the particle. In fact, let us call this function

\[ U(R) = c^2 - c^2 r_S \frac{p_t^2}{R^2} + \frac{p_\varphi^2}{R^2} - \frac{p_t^2 r_S}{R^3}. \]

We may notice that there are terms proportional to \( 1/R^3 \), which we would not expect from the Newtonian gravitation.

This equation (5.3) holds a lot of physical significance. We have the extra term \( 1/R^3 \) in the potential, and we have the relationship between the potential energy, the energy and the radial velocity,

\[ \left( \frac{dR}{d\tau} \right)^2 = c^2 p_t^2 - U. \quad (5.4) \]

\(^1\)In fact, the energy at infinity per unit mass.
Now, clearly \( U(R) = c^2 p_r^2 \) gives us a radial velocity of zero, a point where the radial movement changes direction. Because the derivative of the potential is useful, we note now that it is:

\[
\frac{dU}{dR} = \frac{c^2 r_s R^2 - 2p_r^2 R + 3p_r^2 r_s}{R^4}.
\]

**Crash orbits** \((p_r^2 < 3c^2 r_s^2)\)

Looking at the derivative of \( U \) again, we see that there are no turning points for the radial acceleration if \( p_r^2 < 3c^2 r_s^2 \). These are called crash orbits. It is clear that as \( R \to \infty \), the potential energy approaches \( c^2 \). Ingoing particles clearly crash into the body. Escaping particles clearly cannot overcome the potential energy barrier if \( p_r^2 < 1 \), and they crash, but if this is not the case, they escape to infinity.

**Bound orbits** \((p_r^2 > 3c^2 r_s^2)\)

A bound orbit is a curve that never meets the body or goes to infinity, but has large angular momentum. This is clearly the case for planets and stars and similar celestial bodies. If we study \( dU/dR \) once again, we find that there are turning points for the radial acceleration if \( p_r^2 > 3c^2 r_s^2 \). In fact, there are two:

\[
R^+ = \frac{p_r^2}{c^2 r_s} + \sqrt{\frac{L^4 - 3L^2 r_s^2 c^2}{c^4 r_s^4}},
\]

\[
R^- = \frac{p_r^2}{c^2 r_s} - \sqrt{\frac{L^4 - 3L^2 r_s^2 c^2}{c^4 r_s^4}}.
\]

The orbiting particle is in a potential well around the first \((R^+)\) of these points. It oscillates between the two points \( R \) such that \( U(R) = c^2 p_r^2 \). See figure 5.1 for an energy plot of the potential energy for a bound particle with energy \( c^2 p_r^2 < \max(U) \).

**Figure 5.1** Energy plot for the bound orbit of particles with \( p_r^2 > 3c^2 r_s^2 \), where the particle oscillates between two radii \( R \) such that \( U(R) = c^2 p_r^2 \).

### 5.3 Schwarzschild black holes

The careful reader has probably already identified our indiscriminate use of the Schwarzschild metric, equations (5.1) in Theorem 4.8 and (5.2). Inspired by the
discussions of Schwarzschild black holes in [11] and [16], we try to explain these problems. First, it is clear that \( R = 0 \) is not allowed. However, this is trivial and does not in fact change anything.

Second, we run into trouble when trying to move at \( R < R^* \), where \( R^* \) is the boundary of the body in question. However, this is rarely a problem. Third, however, we notice that at \( R = r_S \) the line element breaks down.

This leads us to break the manifold \( \mathcal{M} \) into two regions:

\[
\mathcal{M}_B = \mathbb{R} \times \{ R : R < r_S \} \times S^2 \quad \text{and} \quad \mathcal{M}_E = \mathbb{R} \times \{ R : r_S < R < \infty \} \times S^2.
\]

Normally, for a star \( R^* > r_S \), and we only work with the Schwarzschild exterior spacetime \( \mathcal{M}_E \), and everything in this chapter so far needs no further mathematical motivation.

However, if we look at an object with \( R^* < r_S \), it is a black hole. In fact, if \( R^* < r_S \), then \( R^* = 0 \), and the mass becomes a point mass, a gravitational singularity. Most of the time, we handle the black hole by ignoring this discontinuity, and using \( \mathcal{M} = \mathcal{M}_E \cup \mathcal{M}_B \).

That is not to say that we do not find the geometrical properties of the black hole interesting.

Looking at the motion of a particle towards a black hole, we notice that the time dilation \( dt = d\tau/(1 - r_S/R)^{1/2} \to \infty \) as \( R \to r_S \). This means that it is not physically possible to observe a particle entering a black hole, while of course it does not take infinite time in the reference frame of the particle itself.

Furthermore, for \( R < r_S \), it is clear that \( 1 - r_S/R \) is negative, and \( t \) becomes a spacelike coordinate while \( R \) becomes the timelike coordinate.
Bibliography


